

# A Trifecta of Modelling Tools: A Bayesian Binary Black Hole Model Selection combining Population Synthesis and Galaxy Formation Models

Liana Rauf<sup>1,2,3\*</sup>, Cullan Howlett<sup>1,2</sup>, Simon Stevenson<sup>2,4</sup>, Jeff Riley<sup>2,5</sup>, Reinhold Willcox<sup>6</sup>

<sup>1</sup>*School of Mathematics and Physics, The University of Queensland, Brisbane, QLD 4072, Australia*

<sup>2</sup>*OzGrav: The ARC Centre of Excellence for Gravitational Wave Discovery, Australia*

<sup>3</sup>*The Research School of Astronomy and Astrophysics, Australian National University, Stromlo, ACT 2601, Australia*

<sup>4</sup>*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Mail H39, PO Box 218, VIC 3122, Australia*

<sup>5</sup>*School of Physics and Astronomy, Monash University, Clayton VIC 3800, Australia*

<sup>6</sup>*Institute of Astronomy, KU Leuven, Celestijnaan 200D, 3001 Leuven, Belgium*

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

Gravitational waves (GWs) have revealed surprising properties of binary black hole (BBH) populations, but there is still mystery surrounding how these compact objects evolve. We apply Bayesian inference and an efficient method to calculate the BBH merger rates in the SHARK host galaxies, to determine the combination of COMPAS parameters that outputs a population most like the GW sources from the LVK transient catalogue. For our COMPAS models, we calculate the likelihood with and without the dependence on the predicted number of BBH merger events. We find strong correlations between hyper-parameters governing the specific angular momentum (AM) of mass lost during mass transfer, the mass-loss rates of Wolf–Rayet stars via winds and the chemically homogeneous evolution (CHE) formation channel. We conclude that analysing the marginalised *and* unmarginalised likelihood is a good indicator of whether the population parameters distribution and number of observed events reflect the LVK data. In doing so, we see that the majority of the models preferred in terms of the population-level parameters of the BBHs greatly overpredict the number of events we should have observed to date. Looking at the smaller number of models which perform well with both likelihoods, we find that those with no CHE, AM loss occurring closer to the donor during the first mass-transfer event, and/or higher rates of mass-loss from Wolf–Rayet winds are generally preferred by current data. We find these conclusions to be robust to our choice of selection criteria.

**Key words:** gravitational waves – black hole physics – stars: evolution – galaxies: star formation – methods: analytical – methods: numerical – methods: statistical

## 1 INTRODUCTION

The detection of gravitational waves (GWs) have prompted the astrophysical community to revise the models for binary black hole (BBH) formation and evolution. There are currently two main formation channels proposed:

(i) Isolated: The formation of merging BBHs through the evolution of isolated binaries. These stars are gravitationally bound since birth and are not perturbed by other stars or compact objects (Mapelli & Giacobbo 2018; Bavera et al. 2021; Santoliquido et al. 2021; Broekgaarden et al. 2021, 2022b; Olejak et al. 2022).

(ii) Dynamical: BBH mergers occurring in dense environments via three-body encounters followed by energy exchange and hardening of the binary. Some examples of dense environments are young star, nuclear or globular clusters, or in accretion disks of the central supermassive BHs in galaxies (Rodríguez et al. 2019; Yang et al. 2019; Arca Sedda et al. 2020; Zevin et al. 2021b).

GWs can be used to gain insight into the poorly constrained aspects of massive stellar binary evolution and help in distinguishing between

binary formation channels. The detailed origins and formation channels for binary systems are still unknown, particularly what fraction of these systems come from isolated, which is the focus of this paper, and dynamical channels.

There are also aspects of stellar evolution that are poorly understood, such as the common envelope phase, mass transfer efficiency and AM loss. This information is critical for understanding the population of BBHs observed by LIGO, Virgo and KAGRA (LVK), including their mass and redshift distributions, and to inform us on what is required for future detectors (Mapelli 2019, 2020; Mandel & Broekgaarden 2022; Arca Sedda et al. 2023). Combining the data from GWTC-1 (Abbott et al. 2019a), GWTC-2 (Abbott et al. 2021a) and GWTC-3 (Abbott et al. 2023b), there is a total of 90 confident ( $p_{\text{astro}} > 0.5$ ; see Section 4.1 and Eq. 20) GW event candidates observed to date, including 85 BBHs, 2 binary neutron stars (BNSs), and 3 black-hole-neutron-stars (BHNSs). This growing population of GWs, including those to be released from the ongoing fourth LVK observing run (O4), provides key information on the source masses, spins and local merger rates (Abbott et al. 2023a) of binary compact objects, which can be used to answer some of these questions.

However, there are many formation pathways that could lead to the same individual source properties. Simulations of ensembles of

\* E-mail: liana.rauf@anu.edu.au

events, produced by varying the physical assumptions, can provide better insight into how these binary systems evolve and merge over time. When combined with the full set of data observed to date, they can potentially break the degeneracy between uncertainties in stellar evolution parameters and BBH populations. They also enable us to produce merger rates that can be compared to observations. [Mandel & Broekgaarden \(2021\)](#) provide a comprehensive review of the local merger rates for the isolated and dynamical formation channels. The studies in the review combine population synthesis codes with different assumptions for the metallicity dependent star formation rate (SFR) density, which plays a crucial role in the evolution of massive binaries ([Neijssel et al. 2019](#); [Chruslińska 2024](#)). [Broekgaarden et al. \(2022a\)](#) utilises COMPAS ([Stevenson et al. 2017](#); [Vigna-Gómez et al. 2018](#); [Team COMPAS: Riley, J. et al. 2022](#)) to explore the impact of stellar evolution parameters and metallicity-specific star formation,  $\mathcal{S}(z, Z)$ , on the merger rate, chirp mass and coalescence time distributions. They find that uncertainties in the binary evolution models only led to variations in the BBH merger rate of up to a factor 2, with the BBH merger rates most sensitive to  $\mathcal{S}(z, Z)$ . [Santoliquido et al. \(2021\)](#) and [Neijssel et al. \(2019\)](#) also come to similar conclusions. [Mennekens & Vanbeveren \(2014\)](#); [Santoliquido et al. \(2021\)](#); [Chu et al. \(2022\)](#) find variations in the merger rate over several orders of magnitude. Many of these studies choose to explore the variations in binary evolution models over a grid of metallicities, and generate populations for each grid point. Some studies have shown that population III stars can also contribute to the total BBH merger rate density ([Santoliquido et al. 2024](#); [Tanikawa 2024](#)). These studies emphasise the need to further study the assumptions in binary evolution models to understand the impact on the merger rate, specifically the *combined* impact of model variations ([Stevenson & Clarke 2022](#)).

There have also been various studies proposing methods for constraining different binary evolution models. [Wong et al. \(2023\)](#) applies “backward” modelling of each GW event to the progenitors to avoid fixed astrophysical assumptions, by allowing variations in the hyper-parameters. However, they note that their methods become less efficient as the dimensionality of the parameter space increases. [Barrett et al. \(2018\)](#) uses Fisher information matrices to quantitatively analyse how the observed BBH population will constrain the COMPAS models, given the chirp mass distribution and accounting for the metallicity-specific cosmic SFR, selection effects and measurement uncertainty. [Riley & Mandel \(2023\)](#) uses artificial neural networks to generate an interpolant to calculate the merger detection rates for double compact objects for varying prescriptions of the metallicity-specific SFR. This is used to create GW populations for each model, where the chirp mass and redshift distribution are compared using Bayesian analysis to the LVK data. [Delfavero et al. \(2023\)](#) calculates a joint likelihood for binary evolution models. They fit normal distribution fits to the likelihood grids and implement Gaussian process regression to interpolate between models, the detected number of events and the likelihood. This is also used to quantify the systematic uncertainty in the parameters. However, they emphasise that the inference on a limited subspace of models could be biased. These studies highlight the high computational costs for multi-dimensional parameter studies. However, as more observations are made, the uncertain astrophysics will produce correlated impacts in the binary population, so there is a crucial need for more parameters covering a large range in model inference. Realistically, all formation channels (within the isolated and dynamical channels) must be considered to obtain a parameter space representative of the true BBH population ([Zevin et al. 2021a](#); [Cheng et al. 2023a](#)).

To build on this body of work, we develop here a new method of combining a more accurate, non-parametric star-formation and

metallicity history of the Universe with a wider range of different massive binary evolution models. We then use these models alongside data from GWTC-1 ([Abbott et al. 2019a](#)), GWTC-2 ([Abbott et al. 2021a](#)) and GWTC-3 ([Abbott et al. 2023b](#)) to identify the most likely avenues for generating binary compact objects, and to rule out those that are unable to produce the distributions of objects seen to date. We do this by improving on [Rauf et al. \(2023\)](#), where we combined runs from the COMPAS rapid population synthesis code, for formation and evolution of BBH formed through isolated binary evolution, with a SFR and metallicity history for the observed host galaxies predicted by the semi-analytic galaxy evolution code SHARK ([Lagos et al. 2018, 2019](#)). In our original work, we predicted (among other things) the BBH merger rate as a function of redshift for four COMPAS runs, where we varied the remnant mass prescription. We found the merger rate ranged over several magnitudes and we concluded that the [Mandel & Müller \(2020\)](#) remnant mass prescription was the best fit out of the four models to the local merger rate currently predicted using the BBH mergers from GWTC-3. However, we also found that cross-matching between the SHARK galaxies and the COMPAS population was prohibitively time-consuming for exploring larger portions of our parameter space. Since we explored only rates, we were also not able to confidently rule out any models that do not align with the GW population-level parameters, and did not test the impact of varying sub-parameters of these models such as those covering other aspects of binary evolution rather than just the mass of the remnant compact objects.

So, in this paper, we substantially extend on our work in [Rauf et al. \(2023\)](#) and explore an alternative and efficient method of predicting the BBH merger rate and population level parameters such as the binary masses and redshift distributions. We do so by rescaling a high-fidelity single run of our original simulation, varying some of the most relevant COMPAS input parameters. This allows us to test in this work  $> 300$  different models for population synthesis, varying ingredients/model assumptions such as the common envelope efficiency, mass transfer accretion efficiency, mass-loss rates of Wolf–Rayet stars, chemically homogeneous evolution model and specific AM loss. For each of these we evaluate the ability of the model to reproduce the observed distribution of black hole masses and redshifts, using a Bayesian framework to account for selection effects. We also explore the impact on our findings with and without including the merger rate as a fitting parameter (i.e., including a comparison of the model-predicted and observed merger rates in evaluating our model posteriors). Compared to previous studies this work also has the improvement of incorporating a basic estimate of sampling error in the simulated model merger rates and posteriors. This allows us to set a threshold for what we consider to be tangible differences in model performance, and also demonstrates more generally the importance of further investigation into the uncertainties in the typical ‘simulation-based inference’ used in GW population studies.

This paper is structured as follows. In Section 2, we summarise recent progress in our understanding of various massive binary evolution models and parameters, motivating the range of models we explore in this work. In Section 3, we show that starting with a fiducial model, where the merger rate in a set of simulated galaxies has been evaluated by carefully cross-matching between COMPAS and SHARK, we can rescale the simulation to produce the time and metallicity dependent GW merger rate for another model where the COMPAS parameters have been altered. In Section 4, we implement Bayesian inference to determine which of these COMPAS runs output a BBH population that best matches the current observed population. We do so by comparing the chirp mass, mass ratio and redshift distributions

from each model with the LVK posterior sample distributions. We then validate the total number of theoretical and detected GW events expected for the models and the models posteriors by comparing to the values obtained using the four ‘brute-force’ evaluations from Rauf et al. (2023). We repeat the method for a wide range of new COMPAS runs in Section 5, identifying the best and worst fit models to the data and studying the impact of different selection effects and analysis choices on the conclusions we can draw about stellar evolution. We conclude in Section 6. In this work, we use the SHARK cosmology  $H_0, \Omega_m = \{67.8, 0.308\}$  in flat  $\Lambda$ CDM where necessary.

## 2 COMPAS: OVERVIEW OF MODELS

In the following section, we detail the physics and the implementation in COMPAS of the common envelope (CE), mass transfer, mass loss via winds, angular momentum (AM) loss and formation channels such as the isolated and chemically homogeneous evolution. We review the literature on the assumptions and parameters associated with these processes that can impact the evolution of the binary stars and, hence, the number and source properties of GWs for each COMPAS model. We conclude with a summary of the big questions surrounding these prescriptions and their limitations in the context of BBH population modelling. Note that in this section, we refer to the more massive star as the donor and the lighter star as the accretor star.

### 2.1 Isolated formation channels

In this paper, we focus on BBHs in the isolated formation channel. This is the most commonly studied channel, but see Fig 1 in Han et al. (2020) for possible binary formation channels. It involves massive stars formed on the zero-age-main-sequence (ZAMS). We describe, one of the formation channels, the classical isolated binary evolution via a CE phase as follows:

- (i) The primary star in the binary reaches the end of main sequence (MS) evolution. After hydrogen burning in the core ends, the core contracts and causes the outer stellar envelope to expand past the Roche lobe (due to the negative heat capacity of the star (Lynden-Bell & Wood 1968)) and transfer mass to the secondary star.
- (ii) Mass transfer from the primary stops and leaves a Wolf-Rayet star.<sup>1</sup> This then reaches the end of its life, undergoes a supernova and forms a BH. The system is now an X-ray binary.
- (iii) The secondary evolves and transfers mass back to the primary BH. This process is dynamically unstable and forms a CE, which encompasses the compact object.
- (iv) Orbital energy released due to the compact object rapidly in-spiralling can remove the CE.
- (v) The secondary star evolves and eventually also forms a BH.
- (vi) The BBH system slowly in-spirals due to orbital energy loss via GWs.

However, there are many stages within this formation channel where the binary may not survive, or the events can subtly differ from the above prescription leading to dramatically different outcomes. The system may not survive the first supernova explosion or the CE may be not be ejected resulting in the compact object merging with the secondary core before a BBH is formed (Schroder et al. 2020; Grichener 2023). Alternatively, the second mass transfer could be

<sup>1</sup> This formation scenario for Wolf-Rayet stars is disputed by Shenar et al. (2020) due to lack of observations of massive stars at low metallicity.

stable and avoid the CE (e.g., van den Heuvel 1976; van den Heuvel et al. 2017; Neijssel et al. 2019; van Son et al. 2022a; Gallegos-Garcia et al. 2021; Marchant et al. 2021). The wind mass-loss rates and remnant prescription can also impact the orbital separation, as well as the final masses. There are various circumstances and factors that can affect the separation of BBH systems, which are strongly correlated with the delay time and the merger rate. On one hand, this makes GW sources unique probes of cosmic history (Neijssel et al. 2019; Mapelli 2021; Mandel & Farmer 2022; Costa et al. 2023), but on the other, leads to difficulties in interpreting or identifying the most important/critical points in the evolution of the system.

As an example, Stevenson et al. (2017) used COMPAS to show that the first three GW events detected by LIGO: GW150914, GW151226 and GW151012, can be explained by the classical isolated channel with mass transfer and CE phase, and characterised their progenitor metallicity and masses. However, Dorozzmai & Toonen (2022) find that the stable channel, which consists of two stable mass transfer phases, dominates the GW sources. The formation of GW sources via a radiative CE channel is not possible for low mass transfer efficiency ( $f_{MT} \leq 0.3$ , see Eq. 1 in Section 2.1.1) due to the mass ratio being below the critical mass ratio ( $q_{crit} = M_d/M_a$ , where  $M_a$  is the accretor mass and  $M_d$  is the donor mass) for instability. Picco et al. (2023) similarly find that stable mass transfer, where mass and AM is lost through the L2 point, is a possible formation channel for BBHs.

Comparing the two channels, van Son et al. (2022a) find that the CE channel tends to dominate BBH mergers with primary masses  $< 30M_\odot$  in binaries with short delay times ( $< 1$ Gyr). They also find that the CE phase is rare for massive stars due to heavy mass loss preventing them from filling their Roche lobe. As such, the stable mass transfer channel, which widens the binary orbit as mass is lost from the system during secondary mass transfer, forms primary BH masses  $> 30M_\odot$  and systems with long delay times ( $> 1$ Gyr). They also find that the CE or stable mass transfer channel dominates the mass distribution depending on redshift, which is in line with Neijssel et al. (2019) finding redshift dependence in the mass distribution. van Son et al. (2022b) studies the stable mass transfer channel only and the impact of  $f_{MT}$  on the local BBH merger rate and the primary BH mass distribution. For  $f_{MT} = 0.75$  and  $f_{MT} = 1$ , there are no BBHs formed through the stable mass transfer channel. The merger rate drops from  $141 \text{ Gpc}^{-3}\text{yr}^{-1}$  with  $f_{MT} = 0.25$  to  $16 \text{ Gpc}^{-3}\text{yr}^{-1}$  with  $f_{MT} = 0.5$ , which implies the stable mass transfer contributes strongly to the merger rate for  $f_{MT} < 0.5$ . The low mass end of the primary BH mass distribution is justified by the stable mass transfer channel — as  $f_{MT}$  increases, the minimum primary BH mass increases. Overall, the relative contribution of both channels is uncertain. The CE channel in particular may be over-represented in population synthesis models (see Section 2.1.1 for relevant references).

#### 2.1.1 Common envelope evolution ( $\alpha_{CE}$ ) and Mass transfer efficiency ( $f_{MT}$ )

Given the above, in the formation of isolated binary stars, uncertainties remain in the CE phase and mass transfer process, which play a key role in the final component and orbital properties.

The mass transfer efficiency is the fraction of mass lost by the donor star and accreted by the companion star. Team COMPAS: Riley, J. et al. (2022) uses a simplified approach for the mass transfer. They parameterise the fraction of mass lost by the donor to the accretor with  $f_{MT}$ ,

$$\dot{M}_a = -f_{MT}\dot{M}_d, \quad (1)$$

with  $0 \leq f_{\text{MT}} \leq 1$  and  $\dot{M}_a$  and  $\dot{M}_d$  are the mass loss rates for the accretor and donor respectively. This fraction depends on the properties of the binary prior to mass transfer and varies over time. We use a fixed fraction prescription in COMPAS, so the fraction of mass accreted is assumed to be constant over the mass transfer phase.  $f_{\text{MT}} = 1$  corresponds to fully conservative mass transfer, so all the mass lost from the donor is accreted.  $f_{\text{MT}} < 1$  is non-conservative and  $f_{\text{MT}} = 0$  is fully non-conservative. We choose to explore the full range of possibilities in COMPAS. We also explore the default setting, where the mass transfer efficiency is determined by the ratio of the Kelvin–Helmholtz thermal timescale of the two stars,

$$\tau_{\text{KH}} = \frac{E_{\text{int}}}{L}. \quad (2)$$

This is the time taken for the star’s internal energy,  $E_{\text{int}}$ , to be radiated at its current luminosity,  $L$ . This ratio translates to a fixed mass ratio,  $0 \leq f_{\text{MT}} \leq 1$ , but depends in the initial mass ratio and orbital separation (Schneider et al. 2015). The maximum accretion rate onto a stellar accretor is,

$$\dot{M}_a = \frac{CM_a}{\tau_{\text{KH},a}}, \quad (3)$$

where the default factor  $C = 10$  is assumed to account for the expansion of the star due to mass loss (Paczynski & Sienkiewicz 1972; Neo et al. 1977; Hurley et al. 2002; Schneider et al. 2015). This is calculated for every mass transfer event and is non-constant. Accretion onto the compact objects, such as black holes and neutron stars, is assumed to be Eddington limited such that,

$$\dot{M}_a = \dot{M}_{\text{edd}} = 1.5 \times 10^{-8} \frac{R_*}{10\text{km}} \frac{M_\odot}{\text{yr}}, \quad (4)$$

where the black hole radius,  $R_*$ , is the Schwarzschild radius. The Eddington accretion limit occurs when the BH accretion disk reaches a maximum luminosity beyond which radiation pressure will overcome gravity, and the infall of material outside of the object will be halted. van Son et al. (2020) show that it is possible to break this limit via stable mass transfer or CE accretion, known as super-Eddington accretion. In the unstable mass transfer scenario, for simplicity, the CE phase is parameterised by COMPAS in the context of energy balance with the  $\alpha_{\text{CE}}$  formalism (Xu & Li 2010) (see also Ivanova et al. 2013, and references therein). The CE efficiency,  $\alpha_{\text{CE}}$ , describes the efficiency of converting the orbital energy into kinetic energy to unbind the CE. This unbinding energy is defined as,

$$E_{\text{bind}} = \alpha_{\text{CE}} \left( \frac{GM_d M_a}{2a_i} - \frac{GM_{\text{d,core}} M_{\text{a,core}}}{2a_f} \right), \quad (5)$$

where  $G$  is the gravitational constant,  $M_d$  is the mass of the pre-CE donor,  $M_a$  is the mass of the accretor,  $M_{\text{d,core}}$  is the mass of the donor’s core,  $a_i$  and  $a_f$  are the pre- and post-CE orbital separations respectively. Increasing  $\alpha_{\text{CE}}$  increases the unbinding energy, so the CE can be ejected more efficiently at wider binary separations. Too small values of  $\alpha_{\text{CE}}$  would require the binary to shrink in order to obtain enough unbinding energy (van den Heuvel 1976; Tutukov & Yungelson 1979; Webbink 1984; Ivanova et al. 2013; Marchant et al. 2021; Hirai & Mandel 2022; Stevenson & Clarke 2022). de Kool (1990) proposed an alternative version to the  $\alpha$ -formalism for the binding energy, introducing a new parameter,  $\lambda$ , to characterise the central concentration of the donor’s envelope,

$$E_{\text{bind}} = -\frac{GM_d M_{\text{env}}}{\lambda a_i r_L}, \quad (6)$$

where  $M_{\text{env}} = M_d - M_{\text{d,core}}$  is the mass of the envelope and

$r_L = R_L/a_i$  is the ratio of the Roche lobe radius and pre-CE orbital separation. The post-CE orbital separation is determined by substituting Eq. 6 into Eq. 5,

$$\frac{a_f}{a_i} = \frac{M_{\text{d,core}} M_a}{M_d} \frac{1}{M_a + 2M_{\text{env}}/\alpha_{\text{CE}} \lambda r_L}. \quad (7)$$

Due to lack of understanding and convenience in population synthesis calculations,  $\alpha_{\text{CE}}$  and  $\lambda$  are set to constants, but *should* be varied depending the stellar and binary parameters (Xu & Li 2010). For  $0 \leq \alpha_{\text{CE}} \leq 1$ , the only source of energy to unbind the CE is the orbital energy. However, various hydrodynamical simulations over the years have shown that the orbital energy alone might not be sufficient for envelope removal (Ivanova et al. 2013; Röpke & De Marco 2023). Additional energy sources are required ( $\alpha_{\text{CE}} > 1$ ) such as recombination energy from evolved giant star envelopes (Ivanova et al. 2015; Lau et al. 2022) or the giant stars can launch jets as they accrete mass via an accretion disk (Ohlmann et al. 2016; Moreno Méndez et al. 2017). In this paper, we explore  $0.1 \leq \alpha_{\text{CE}} \leq 100$ . Although our upper limit seems extreme, it should be noted that Hirai & Mandel (2022) implements a two-step CE formalism which is equivalent to  $\alpha_{\text{CE}} > 10$  in most cases (see their Fig. 5). This emphasises the need for additional sources of energy, but it may also suggest that the current  $\alpha_{\text{CE}}$ -formalism in COMPAS is wrong. As such, we choose to include  $\alpha_{\text{CE}} = 100$  as an exploratory parameter. See further discussion of this in Section 2.4.

Progenitors of GW sources provide an alternative path for comparison between models and observations. Zuo & Li (2014) uses the binary population synthesis code by Hurley et al. (2000) and Hurley et al. (2002) to study high-mass X-ray binaries (HMXB) and the correlation between their apparent luminosity and displacement from star clusters. A high value of  $\alpha_{\text{CE}} = 0.8 - 1.0$  is preferred when comparing to observations by Kaaret et al. (2004). However, the ambiguity in the definition of the donor core mass boundary introduces a large uncertainty (almost two orders of magnitude), which translates directly to the expected value of  $\alpha_{\text{CE}}$ . Romero-Shaw et al. (2023) uses COMPAS to study black hole high-mass X-ray binaries (BH-HMXB). They find for the case where  $f_{\text{MT}} \leq 0.1$  (and  $f_{\text{WR}} = 0.1$ , explained in Section 2.1.2), that 8% of BH-HMXBs end up as BBHs or BHNSs. de Mink et al. (2007) attempts to fit the mass transfer efficiency fraction for massive close binaries. They find a wide spread in  $f_{\text{MT}}$  and conclude that a single value is not valid. Therefore, more electromagnetic observations are required to constrain modelling degeneracies of massive binaries.

There have been many studies on the mass transfer and CE phase. Many of these have explored how varying these parameters affect the BBH merger rate, the BH mass distribution range and which binary formation channels dominate these BBH mergers (van Son et al. 2020, 2022a,b). Many of these studies also find that additional energy sources are necessary to produce BBH mergers and other double compact objects (DCO) (Santoliquido et al. 2021; Zevin et al. 2021a; Marchant et al. 2021; Grichener 2023; Deng et al. 2024). To infer the formation channels of BBH mergers, many studies further explore the connection between the models and data (Cheng et al. 2023b). Bouffanais et al. (2021) find the preferred parameter range for the CE and mass transfer efficiency is  $[4, 7]$  and  $[0.7, 0.8]$  respectively. They also showed that increasing the CE efficiency increases the uncertainty on the local BBH merger rate density. Fragos et al. (2019) find that the final orbital separation for their BNS’s to merger within a Hubble time translates to  $\alpha_{\text{CE}} \sim 5$ . Broekgaarden & Berger (2021) and García et al. (2021) also find a preference for  $\alpha_{\text{CE}} > 1$  for producing merger rates compared to the observed rates. However studies such as Bavera et al. (2021), Klencki et al. (2021) and Cui &

Li (2023) find there is a preference for lower  $\alpha_{\text{CE}}$  ( $0.2 < \alpha_{\text{CE}} < 1$ ,  $\alpha_{\text{CE}} \leq 0.7$  and  $\alpha_{\text{CE}} = 0.5, 1$  respectively). Klencki et al. (2021) further questions the validity of the CE channel for producing massive BHs. They find that successful ejection of the CE is possible if the donor is a massive convective-envelope giant.

We briefly mention the various studies that explore the uncertainties in parameters. But another uncertainty to note is in how these parameters are implemented in population synthesis codes. Gallegos-Garcia et al. (2021) highlights the uncertainty in stellar evolution tracks by comparing the mass transfer and CE treatments in the detailed stellar evolution code, MESA (Paxton et al. 2011a, 2013, 2015, 2018, 2019), and rapid binary population synthesis code, COSMIC (Breivik et al. 2020a). The COSMIC models primarily form BBH mergers through the CE phase, while the stable mass transfer formation channel is preferred with MESA. Increasing  $\alpha_{\text{CE}}$  also increases the variation in the BBH merger rates between the models, due to variations in merger times. Agrawal et al. (2022a) and Romagnolo et al. (2023) have compared population synthesis codes to explore the uncertainty in radial expansion of massive stars and the variation in the remnant mass.

### 2.1.2 Wolf-Rayet Winds ( $f_{\text{WR}}$ )

Stellar winds are outflows of gas from the atmosphere of a star. Wolf-Rayet stars are evolved stars with high mass loss rates, where the hydrogen envelope of the stars are removed. To account for uncertainties in the mass-loss rates of Wolf-Rayet stars, in COMPAS, the Wolf-Rayet multiplier,  $f_{\text{WR}}$ , is used to parameterise the mass loss rate (Belczynski et al. 2010),

$$\frac{dM}{dt} = f_{\text{WR}} \times 10^{-13} L^{1.5} \left(\frac{Z}{Z_{\odot}}\right)^m M_{\odot} \text{yr}^{-1}, \quad (8)$$

where  $M$  is the mass of the helium star,  $L$  is the luminosity,  $Z$  is the metallicity,  $Z_{\odot} = 0.014$  (Asplund et al. 2009) and  $m = 0.86$  (Vink & de Koter 2005). Wolf-Rayet stars have been studied to understand their explosion mechanisms and determine if they are progenitors of gamma ray bursts (Detmers et al. 2008). More recently, Wolf-Rayet stars have been used to study GW sources (Vink & Harries 2017; Bogomazov et al. 2018). Decreasing  $f_{\text{WR}}$  leads to an increase in the BH masses (Dorozsmai & Toonen 2022; Romero-Shaw et al. 2023). In COMPAS, we choose the Wolf-Rayet multiplier range provided in Stevenson & Clarke (2022), which accounts for the range of uncertainties in Wolf-Rayet mass-loss rates.

HMXBs are usually ideal candidates to study and constrain the effects of stellar mass loss in interacting binaries, as radiation driven winds are expected for giant stars. Schröder et al. (2021) finds that the orbital evolution of HMXBs depends on the wind velocity and mass ratio. Slow winds and high mass objects results in the binary shrinking, due to drag forces. In the case where there is no drag, the binary separation widens. GWs can provide an indirect approach to study the separation of binaries, as their progenitors include HMXBs. However, Liotine et al. (2023) find that no currently observed HMXB is predicted to form a merging BBH with high probability. They use COSMIC (Breivik et al. 2020b) to conclude that the BHs in the detectable HMXBs are not likely to have masses  $> 35M_{\odot}$ , the detectable HMXBs will not merge as BBHs within the Hubble time and the distribution of redshifts, masses and metallicities of detectable HMXBs and BBHs are different. Fishbach & Kalogera (2022) argue that discrepancies in the BH masses in BBHs and HMXBs are due to GW observational selection effects.

Neijssel et al. (2021) and Romero-Shaw et al. (2023) find a preference for low  $f_{\text{WR}}$  ( $< 0.2$ ) for HMXB systems such as Cygnus

X-1 (Miller-Jones et al. 2021). Stevenson & Clarke (2022) explores the correlated impact of model parameters on the merger rate. They identified a correlation between the Wolf-Rayet factor and metallicity, as increasing both parameters increases mass loss via stellar winds. They find most models that match the current observations have  $f_{\text{WR}} > 1$ . This contradicts the previously mentioned studies due to COMPAS over-predicting the rates and masses of BBHs formed through chemically homogeneous evolution (see Section 2.2).

Riley et al. (2021) studies the CHE binaries as a formation channel for various  $f_{\text{WR}}$  with COMPAS. They find a correlation between the delay time and  $f_{\text{WR}}$ , as increasing mass loss leads to wider binaries. This results in the merger rate decreasing with increasing  $f_{\text{WR}}$ . For  $f_{\text{WR}} = 1$ , the CHE channel contributes  $20 \text{ Gpc}^{-3} \text{ yr}^{-1}$  to the total rate of  $50 \text{ Gpc}^{-3} \text{ yr}^{-1}$ , which is an overestimation compared to GWTC-3. Due to the short delay times, the BBH merger rate peaks at higher redshifts ( $z \approx 3 - 4$ ) compared to the peak SFR.

### 2.1.3 Specific AM loss ( $f_{\gamma}$ )

Mass transfer from a massive donor to a lower-mass accretor adds specific AM to the transferred mass, so conservation of AM requires that binary's orbit shrinks. If mass transfer is non-conservative, angular momentum,  $J_{\text{orb}}$ , is carried away by mass lost from the system. The change in AM is parameterised by  $\gamma$ ,

$$\frac{\dot{J}_{\text{orb}}}{\dot{M}_d} = \gamma \frac{J_{\text{orb}}}{M_d}. \quad (9)$$

The orbital evolution of the binary is given by,

$$\begin{aligned} \frac{\dot{a}}{a} &= -2 \frac{\dot{M}_d}{M_d} \left[ 1 - f_{\text{MT}} \frac{M_d}{M_a} - (1 - f_{\text{MT}}) \left( \gamma + \frac{1}{2} \right) \frac{M_d}{M_d + M_a} \right] \\ &= -2 \frac{\dot{M}_d}{M_d} \left[ 1 - \frac{f_{\text{MT}}}{q} - (1 - f_{\text{MT}}) \left( \gamma + \frac{1}{2} \right) \frac{1}{1+q} \right], \end{aligned} \quad (10)$$

where  $q = M_a/M_d$  is the mass ratio. Willcox et al. (2023) introduces a new parametrisation,  $a_{\gamma}$ ,

$$a_{\gamma} = a_{\text{acc}} + f_{\gamma} (a_{\text{L2}} - a_{\text{acc}}), \quad (11)$$

which is the effective decoupling radius between the accretor,  $a_{\text{acc}}$ , and the L2 point,  $a_{\text{L2}}$ . COMPAS default model assumes isotropic re-emission of matter during non-conservative stable mass transfer ( $f_{\gamma} = 0$ ). In this model, mass lost from the donor is transported near the accretor and ejected via fast isotropic winds. Hydrodynamical simulations indicate that this mass is lost between the accretor and L2 point (MacLeod et al. 2018a,b).  $f_{\gamma}$  determines the point where this AM is lost. If the donor star is growing rapidly with respect to its Roche Lobe, the system will lose mass through the outer Lagrangian point (L2), which causes the orbit to shrink. This corresponds to  $f_{\gamma} = 1$ . AM loss directly from the accretor corresponds to  $f_{\gamma} = 0$  (Costa et al. 2023; Romero-Shaw et al. 2023). We consider  $0 \leq f_{\gamma} \leq 1$ . It is possible for cases outside of this range such as intense feedback radiation from the accretor for mass loss at or near the the inner Lagrangian point, L1, ( $f_{\gamma} < 0$ ) or magnetic fields that keep material in co-rotation beyond the L2 point ( $f_{\gamma} \geq 1$ ) (MacLeod & Loeb 2020; Willcox et al. 2023).

Lu et al. (2023) and Marchant et al. (2021) find that their models predicted high mass transfer rates, which lead to mass loss through the L2 point. MacLeod et al. (2018a); MacLeod & Loeb (2020) argue that the specific AM is ejected between the accretor and the L2 point. Willcox et al. (2023) uses COMPAS to study the stable mass transfer channel. They find that increasing  $f_{\gamma}$  leads to a reduction in the number of systems that undergo stable mass transfer. Increasing  $f_{\text{MT}}$

while decreasing  $f_\gamma$  increases the likelihood of stable mass transfer. For  $f_\gamma \gtrsim 0.7$ , nearly all systems experience CE evolution.

## 2.2 Chemically homogeneous evolution

In the chemically homogeneous evolution (CHE) channel, the stars do not expand. They are metal poor and rapidly rotating in tight binaries. The rotation efficiently transports hydrogen to the core and helium to the envelope. This evolves and forms a Wolf–Rayet star, which contracts rather than expands. Wind-driven mass loss does not widen the binary significantly for low metallicity systems and there is no mass transfer, with the exception of massive over-contact binaries (Marchant et al. 2016). Highly spinning DCOs would produce GWs in the age of the Universe (Nathaniel 2017; Mandel & Farmer 2022; Costa et al. 2023). de Mink et al. (2009, 2010) introduces the concept of CHE to justify short period massive binaries such as the HMXB M33 X-7 (Orosz et al. 2007).

Mandel & de Mink (2016) find that the CHE channel significantly contributes to the BBH merger rate ( $10.5 \pm 0.5 \text{ Gpc}^{-3} \text{ yr}^{-1}$  at redshift 0). These systems typically have a total mass ranging between 50–110  $M_\odot$ . Marchant et al. (2016) explores the parameter space for forming CHE binaries. They find that at low metallicity, BBHs will merge within the age of the Universe due to weaker winds ( $t_{\text{delay, min}} = 0.4 \text{ Gyr}$  for  $Z_\odot/50$ ). There is strong evidence for increasing masses and mergers with decreasing metallicity (Belczynski et al. 2010; Bulik & Belczynski 2010; Kowalska-Leszczynska et al. 2015; Spera et al. 2015; Neijssel et al. 2019), and for masses in the pair instability SN gap (van Son et al. 2020; Winch et al. 2024).<sup>2</sup> If these mergers were detected, they could probe the evolution of massive stars in the early Universe. In addition to calculating the merger rate and mass distribution for the CHE channel, du Buisson et al. (2020) finds that the cosmological merger rate ( $5.8 \text{ Gpc}^{-3} \text{ yr}^{-1}$  at redshift 0) can vary up to 40% for extreme deviations in the metallicity-specific SFR.

In COMPAS, there are two CHE modes, ‘optimistic’ and ‘pessimistic’, which refer to how a CH star evolves on the MS (Riley et al. 2021). In the ‘optimistic’ mode, if the star begins the main sequence with fast enough rotation to evolve as a CHE star, it is assumed the star continues rotating fast enough to remain CH throughout the MS and the spin is not checked. For the CHE mode ‘pessimistic’ the spin is checked at every timestep, and if it has slowed below the CH metallicity-dependent rotational frequency threshold, it is immediately converted to a regular MS star and evolve to the end of the MS (Team COMPAS: Riley, J. et al. 2022). We choose to study both variations in COMPAS, as well as no CHE, to determine if the CHE can be ruled out as a formation channel for GW sources.

## 2.3 Remnant mass prescription

GWs can be used to infer the underlying mass distribution of BBHs, which depends on the metallicity and the type of explosion mechanism. Their formation can be described in three stages; stellar collapse and bounce, convective engine, and post-explosion fallback. Massive stars undergo a series of successful burning stages, building an iron core until electron degeneracy pressure can no longer support this. Electron capture and dissociation of the core elements into alpha particles accelerates the core collapse. This collapse is halted when

<sup>2</sup> Due to the production of electron positron pairs in the core of massive stars, the reduction of pressure followed by gravitational collapse leads to oxygen burning and results in a partial or complete explosion, which leaves behind no compact remnant (Fowler & Hoyle 1964; Barkat et al. 1967).

the core reaches nuclear densities and nuclear forces cause the core to bounce and emit a shock outwards. This results in a mixing of the hot and cold layers in the star, known as convection. The convective engine provides the means to convert potential energy released in the collapse into explosion energy in the form of a supernova (Fryer et al. 2012; Fryer et al. 2021). The commonly accepted supernova explosion mechanism is neutrino driven, where some of the neutrinos emitted during the core collapse are reabsorbed, powering an explosion (Janka et al. 2007; Marek & Janka 2009; Janka 2012, 2017). We choose to explore all remnant mass prescriptions available in the COMPAS version used in this paper, and describe them below.

Fryer et al. (2012) provide analytic prescriptions for estimating the remnant mass in population synthesis codes, based on at the time recent understanding of supernova and gamma-ray burst explosions. In their model/fit, the remnant mass depends on the core, fallback mass and explosion mechanism. They explore two neutrino-driven explosion mechanisms—the rapid convection, which leads to explosion in the 250 ms after bounce, and the delayed convection, which can occur over a longer timescale. Fryer et al. (2022) altered the remnant mass prescription from Fryer et al. (2012) to study the growth rate of convection. They introduce a new parameter to describe the mixing growth time for convective instabilities in their remnant mass models, denoted  $f_{\text{mix}}$ . A fast growth time,  $f_{\text{mix}} = 4.0$ , and a slow growth time,  $f_{\text{mix}} = 0.5$ , correspond to the rapid and delayed models respectively from Fryer et al. (2012). The growth time determined the supernova explosive energy (see Fig. 4 in Fryer 2015). Whether a star undergoes a successful or failed supernova provides clues on the final remnant object. However, there are uncertainties surrounding these mechanisms (Mapelli 2021). GW detection from a core collapse SN can be used to distinguish between explosion mechanisms (Powell et al. 2016, 2023)

Mandel & Müller (2020) develop probabilistic prescriptions for the remnant mass based on results from 3D supernovae simulations in Müller et al. (2016). Instead of a single value prescription, the BH mass branches over a range of core masses, depending on the fallback strength. If the BH is not formed by complete fallback, the mass follows as normal distribution. These models are more realistic as they capture the stochasticity of stellar evolution.

Schneider et al. (2021) also applies the Mandel & Müller (2020) semi-analytic SN prescription, but the resulting remnant mass prescription differs. The model focuses more on the removal of the hydrogen-rich envelope of the progenitor star and the evolution steps before core collapse. They also explore the amount of fallback. Their models overestimate the BH masses by 10% if the BHs are not formed by direct collapse. They do not account for any convective boundary mixing beyond the core helium burning.

Schneider et al. (2023) uses the models in Schneider et al. (2021) and the stellar evolution code MESA (Paxton et al. 2011b) to explore the black hole mass distribution and chirp mass distribution of BBH mergers. The ability of stars to explode in the delayed neutrino-driven SN mechanism is characterised by the compactness of the progenitor cores (O’Connor & Ott 2011; Ertl et al. 2016). This depends on the mass, radius and a nuclear reaction network, and a high compactness indicates unsuccessful explosions and collapse into a BH. Thus, the chirp mass distribution is bimodal and the model predicts a lack of chirp masses in the range 9 – 13  $M_\odot$ . Mass accretion can influence the final outcome of core collapse and final remnant mass. Schneider et al. (2024) studies models with MESA that allow for multiple mergers. They find BH masses in the pair instability SN gap, which can explain events such as GW190521. They note that for a more complete picture of stellar mergers, further detailed multi-

dimensional, hydrodynamic simulations and a systematic exploration of the relevant parameter space is needed.

## 2.4 Summary

We outline the parameters and model assumptions varied in COMPAS in Table 1. We note that the initial conditions and distributions for the binary systems are set to default (except for the metallicity where we use SHARK), as shown in Table A1. Future studies may choose to vary these parameters. As discussed, we are aware of the limitations and approximations made in COMPAS and other population synthesis codes, that may cause discrepancies in the outputs. We summarize these limitations and the key questions below:

(i) The binding energy formalism utilises a fixed  $\alpha_{\text{CE}}$ , but is this a valid approach? The parametrisation of  $\alpha_{\text{CE}}$  is a point of contention, particularly when additional energy sources are available. It is also not clear how these would scale with the orbital separation and mass, so the  $\alpha$ -formalism is not an accurate representation of CE ejection. Hirai & Mandel (2022) proposes a two-step CE process, which depends on the evolutionary stage of the donor and the companion's mass. They find that the predicted post-CE separation is wider compared to the  $\alpha_{\text{CE}}$  formalism. This may pose a problem for BBHs merging within Hubble time. Picker et al. (2024) has recently implemented fits for the convective envelope mass and associated values of  $\lambda$  to be used in conjunction with the two-step CE formalism. They find the effective combined value of  $\alpha_{\text{CE}}\lambda$  can range from  $\sim 10^{-3}$  to  $10^3$  when converted to the  $\alpha$ -formalism.

(ii) Are the masses in COMPAS overestimated and how does this impact the CHE binaries? CHE stars evolve from the hydrogen-ZAMS directly to the helium-ZAMS. When a CHE star is close to the end of the MS, it is already rich in helium, and may start to resemble a helium/ Wolf-Rayet star. However, COMPAS implements the same wind loss prescription throughout the MS, instead of a specific Wolf-Rayet wind prescription with higher mass loss rates. This change is applied to the MESA CHE modelling (Marchant et al. 2016; du Buisson et al. 2020; Marchant et al. 2023).

(iii) Mass loss via radiation-driven winds is still an uncertainty in massive star progenitor models and accurate predictions are necessary for predicting their evolutionary paths. Björklund et al. (2023) derive radiation-driven mass-loss from hot, massive stars depending on their fundamental stellar parameters and implement this in MESA. They find that lower mass-loss rates potentially allow for the creation of high-mass black holes even at higher metallicities compared to previous models.

(iv) At ZAMS masses  $> 50M_{\odot}$ , Hurley et al. (2000) smoothly extrapolate the Pols et al. (1998) model grid up to  $150M_{\odot}$ , which is not well tested outside of the prescribed mass range. We find the contribution of  $M_{\text{ZAMS}} > 50M_{\odot}$  to BBH mergers in COMPAS is non-negligible. This mass range could be significant as LVK detects more GW events in the pair instability mass gap (Winch et al. 2024). Significant improvement on this extrapolation method is far from trivial, as the evolution of high mass binaries is still very uncertain. This is reflected by the discrepancies between population synthesis codes in their stellar tracks (Agrawal et al. 2020, 2022a,b).

(v) There are many uncertainties in the mass transfer efficiency and AM loss, which influence the Roche lobe response to mass loss. Uncertainties in how a donor star responds to mass loss, combined with the Roche lobe response uncertainties contributes to the lack of understanding around the mass transfer stability boundary as a function of donor mass and radius (Ge et al. 2010, 2015, 2020; Temmink et al. 2023).

(vi) Is the AM loss occurring closer to the accretor or the L2 point? And should the implementation of  $f_{\gamma}$  vary for degenerate and non-degenerate accretors? When performing the analysis in this paper, split functionality for AM loss from degenerate and non-degenerate accretors was yet to be implemented in COMPAS. Physically, it is more accurate for  $f_{\gamma}$  to be implemented differently for the primary and secondary star, particular when the accretor is a compact object. It is not clear how this would impact our current results, where we have only considered the AM loss near the non-degenerate object. Gallegos-García et al. (2023) derive an analytic model for mass loss from the binary that is informed by observations of X-ray binary accretion disks and hydrodynamical simulations, which considers AM loss from a disk wind around a compact object accretor. They find their model predicts more AM loss compared to the standard case of isotropic re-emission and the orbital separation of a binary can be considerably affected by mass loss via a disk wind.

(vii) The COMPAS version used in this paper does not consider tidal evolution, but it is believed to play a significant role in the evolution of massive stars (Team COMPAS: Riley, J. et al. 2022). Tidal forces can bring the compact object into the envelope of the secondary which is now a giant star. This implies that COMPAS underestimates the number of BH-giant system mergers, while overestimating the number of BBH mergers (Grichener 2023). When performing the analysis in this paper, the newest version of COMPAS implemented tides (Kapil et al., in prep.).

(viii) COMPAS has two modes; single stellar evolution (SSE) and binary stellar evolution (BSE). The Hurley et al. (2000, 2002) SSE algorithms are implemented in the BSE mode at all evolutionary phases and provides the stellar attributes for both components of the binary. However, the fitting formulae are not well behaved outside of the metallicity range  $Z \in [10^{-4}, 0.03]$ .

(ix) The SSE tracks in COMPAS are based on non-rotating stars and do not account for effects on stellar evolution due to mild rotation. Thus, the modelling of rotational mixing, core overshooting (convection beyond the Schwarzschild criterion) and donor stripping is not accurate. Many studies had investigated the impact of these parameters on stellar structures using MESA and PARSEC (Claret & Torres 2018; Costa et al. 2019; Laplace et al. 2020; Temaj et al. 2024).

## 3 PREDICTING THE GW MERGER RATE FOR DIFFERENT COMPAS MODELS

In this section we demonstrate our method for rescaling the merger rate from a fiducial, carefully cross-matched, simulation of galaxy formation (via SHARK) and binary population synthesis (from COMPAS). This allows us to produce predictions for the rate and properties of an ensemble of GW events for any model without the large computational cost of redoing the cross-matching from scratch. From Rauf et al. (2023), the number of GW events per SHARK galaxy that occur at time  $t_j$  from stars born at time  $t_i$  using COMPAS model  $\alpha$  was written as

$$\frac{dN_{\text{gal}}^{\alpha}}{dN_{\text{gal}}}(Z, t_i, t_j) = \frac{\psi(Z, t_i)}{M_{\text{tot}}^{\alpha}} \Delta t_i \sum_n^{N_{\text{binary}}} \vartheta(t_i + \tau_n^{\alpha}(Z), t_j, \Delta t_j), \quad (12)$$

where

$$\Delta t_j = t_j - t_{j-1}, \quad (13)$$

$$\vartheta(t, t_j, \Delta t_j) = \begin{cases} 1 & \text{if } t_j - \frac{\Delta t_j}{2} < t < t_j + \frac{\Delta t_j}{2}, \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

Parameters	Value	Default
Remnant mass prescription	[MULLERMANDEL, SCHNEIDER2020, FRYER2012, FRYER2022]	MULLERMANDEL
Chemical homogeneous evolution	[NONE, PESSIMISTIC, OPTIMISTIC]	PESSIMISTIC
Mass transfer fraction ( $f_{\text{MT}}$ )	[0.1, 0.2, 0.5, 1]	0.5
Common envelope efficiency ( $\alpha_{\text{CE}}$ )	[0.1, 1, 10, 100]	1.0
L2 point AM loss linear fraction ( $f_{\gamma}$ )	[0, 0.25, 0.5, 0.75, 1]	0.5
Wolf-Rayet winds multiplier ( $f_{\text{WR}}$ )	[0.1, 0.5, 1, 5, 10]	1.0

**Table 1.** COMPAS hyper-parameters explored in this paper. Note that some prescriptions are only used in conjunction with others. The mass transfer fraction is only applied if the mass transfer accretion efficiency prescription is set to FIXED\_FRACTION. The default is THERMAL. Similarly, the AM loss linear fraction is only applied when the mass transfer AM loss prescription is MACLEOD\_LINEAR. The default is ISOTROPIC ( $f_{\gamma} = 0$ ). Refer to Tables C1, C2 and C3 in Appendix C for all relevant prescriptions used for each model.

and  $\psi(Z, t_i)$  is the metallicity ( $Z$ ) dependent SFR in the galaxy when the progenitor stars were born.  $M_{\text{tot}}^{\alpha}$  is a normalisation constant for a given model  $\alpha$ <sup>3</sup> (see Eq 10 in Rauf et al. 2023), which normalises the SFR based on the chosen stellar Initial Mass Function (IMF), in such a way that it cancels out the dependence of the number of GW events on the number of binaries that we simulate  $N_{\text{binary}}$ , and any lower or upper mass cuts we apply when running COMPAS.  $\theta(t, t_j, \Delta t_j)$  is a counter for the number of BBH mergers in SHARK snapshot,  $j$ .  $\tau_n^{\alpha}(Z)$  is the coalescence time of each binary outputted from the model  $\alpha$ . If the merger time,  $t = t_i + \tau_n^{\alpha}(Z)$ , falls in the  $j^{\text{th}}$  bin, a count is added to that bin, otherwise it does not contribute. Overall, this equation encapsulates that the rate consists of two contributions, one from the SHARK host galaxy and the other from population synthesis (in our case from COMPAS).

### 3.1 The effective factor for different models

Cross-matching the star formation history of every galaxy in a simulation with a COMPAS run is computationally expensive. To simplify this, we now demonstrate how we can rescale the rate. If instead of considering GW events on a per galaxy basis we consider them in terms of number per metallicity and time  $t_i$  and  $t_j$  bins, we can write

$$\begin{aligned} \frac{d^3 N_{\text{GW}}^{\alpha}}{dZ dt_i dt_j} &= \frac{d^3 N_{\text{gal}}}{dZ dt_i dt_j} \frac{dN_{\text{GW}}^{\alpha}}{dN_{\text{gal}}} \\ &= \frac{d^3 N_{\text{gal}}}{dZ dt_i dt_j} \frac{\bar{\psi}(Z, t_i)}{M_{\text{tot}}^{\alpha}} \Delta t_i \sum_n \theta(t_i + \tau_n^{\alpha}(Z, t_i), t_j, \Delta t_j). \end{aligned} \quad (15)$$

In the second line we have substituted in Eq. 12 and use  $\bar{\psi}$ ,  $\bar{\theta}$  to acknowledge that this is the average of the star-formation history and Heaviside function over all galaxies that fall into the the bin  $dZ dt_i dt_j$  (i.e., we assume that all galaxies in this bin produce the same number of gravitational waves, which is strictly true only for infinitesimally small bins). We now can introduce a new population synthesis model,  $\beta$ <sup>4</sup>, where the merger rate is unknown. If we have the merger rate from model  $\alpha$ , we can take a ratio of the two models

and rearrange,

$$\begin{aligned} \frac{d^3 N_{\text{GW}}^{\beta}}{dZ dt_i dt_j} &= \frac{M_{\text{tot}}^{\alpha}}{M_{\text{tot}}^{\beta}} \frac{d^3 N_{\text{GW}}^{\alpha}}{dZ dt_i dt_j} \frac{\sum_n \theta(t_i + \tau_n^{\beta}(Z, t_i), t_j, \Delta t_j)}{\sum_n \theta(t_i + \tau_n^{\alpha}(Z, t_i), t_j, \Delta t_j)} \\ &= \frac{M_{\text{tot}}^{\alpha}}{M_{\text{tot}}^{\beta}} \frac{d^3 N_{\text{GW}}^{\alpha}}{dZ dt_i dt_j} f_{\text{eff}}^{\alpha\beta}(Z, t_i, t_j, \Delta t_j). \end{aligned} \quad (16)$$

We hence find that retaining the information from SHARK is not required as the star formation history,  $\bar{\psi}(Z, t_i)$ , and the galaxy distribution are independent of the population synthesis model. In practice we also expect that for most use cases  $\frac{M_{\text{tot}}^{\alpha}}{M_{\text{tot}}^{\beta}} \approx 1$  since changes to the IMF between different literature values primarily affect stars that are too small to become binary black holes or neutron stars, although we retain it here for completeness.<sup>5</sup> The only caveat then is on the free choice of binning used to evaluate what we call the ‘effective factor’  $f_{\text{eff}}^{\alpha\beta}$  between the two models — although the dependence on galaxy simulation has been absorbed into our fiducial  $\frac{d^3 N_{\text{GW}}^{\alpha}}{dZ dt_i dt_j}$ , there is still the implicit assumption that the rate in each galaxy in the bin is the same. We find that this approximation remains valid over a range of metallicity binning choices, even for the quite wide bins we adopt as our default (100 un-evenly spaced bins for  $\log Z \in [-7.0, -0.3]$ ). We show an example distribution of  $f_{\text{eff}}$  in these metallicity and merger time bins in Fig. 1 for a single choice of binary formation redshift of 2.5 (corresponding to  $t_i \approx 2.62\text{Gyr}$ ).

### 3.2 Generating the merger rate

Here we give a qualitative overview of our algorithm for predicting the GW merger rate for our different COMPAS models.

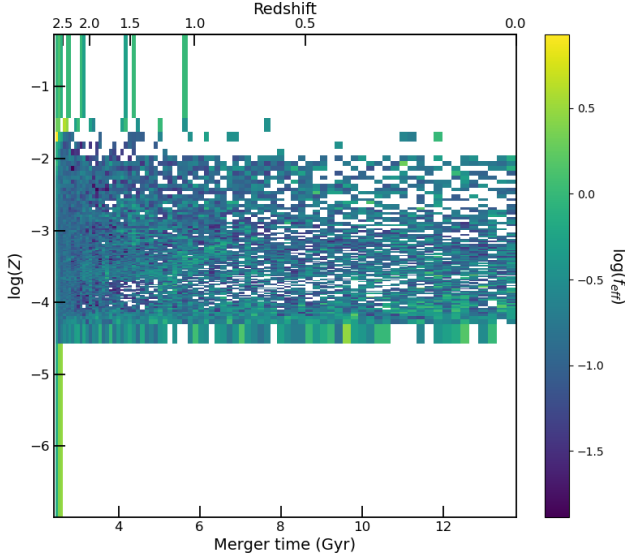
(i) The starting point of our algorithm is a high-fidelity set of GW events simulated from a combination of SHARK and COMPAS model  $\alpha$ . The detailed steps used for generating these are covered in Rauf et al. (2023) and so will not be repeated here. Briefly, this is done by cross-matching the galaxy metallicities to the binary population metallicities outputted from COMPAS. We inject this population into the SHARK galaxies and track their evolution until they merge using Eq. 14.

(ii) From this,  $\frac{d^3 N_{\text{GW}}^{\alpha}}{dZ dt_i dt_j}$  is calculated by taking the fiducial merger rate outputs and binning in terms of metallicity and birth and merger time. For the metallicity we use quantiles to determine the bin edges such that there is an even number of galaxies in each bin. For the

<sup>3</sup>  $\alpha$  refers to some combination of hyper-parameters. In Rauf et al. (2023), it refers to changing the remnant mass prescription (see Table 1).

<sup>4</sup>  $\beta$  now refers to the new COMPAS models in this paper and is some combination of the hyper-parameters in Table 1.

<sup>5</sup> Note this factor also needs to be included if one were to rescale between models where different lower or upper mass limits have been used to draw stars from the IMF, where a different mass ratio distribution for the initial binaries is assumed, or the overall binary fraction is changed



**Figure 1.** An example of the effective factor between two COMPAS models with different remnant mass prescriptions  $\alpha = \text{Fryer et al. (2012)}$  and  $\beta = \text{Mandel \& Müller (2020)}$  in bins of merger time and metallicity for binaries forming at redshift 2.5 (corresponding to  $t_i \approx 2.62\text{Gyr}$ ). In this plot and the rest of our work, an un-equal metallicity binning is chosen such that the number of SHARK galaxies in each metallicity bin in the same. Unshaded bins indicate regions of the parameter space where there are no gravitational waves forming in either one or both of the models.

birth and merger time, we simply use the 170 snapshots from SHARK to define the 170 bins.

(iii) Having established our metallicity, birth time and merger time bins, we take our full list of COMPAS binaries, read in the metallicity and evolution time, which is added to the birth time ( $t_{\text{birth}} + t_{\text{BBH}} + t_{\text{coalescence}}$ ), and add these to the corresponding bins. We apply this to both the fiducial and new model, take the ratio and multiply by the merger rate in (ii) to estimate the new merger rate, shown in Eq. 16.

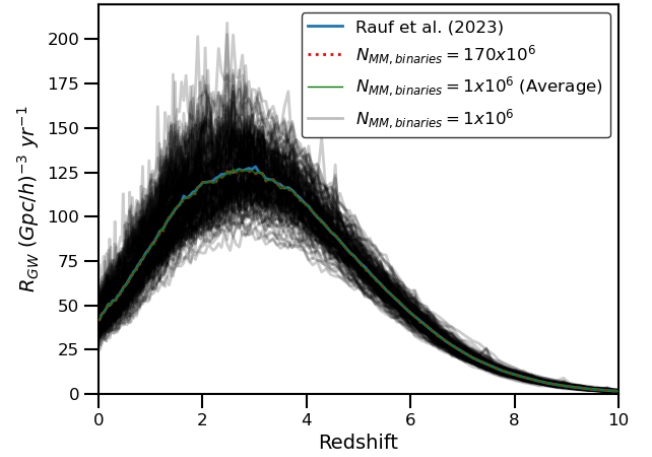
Once we have our estimate for the number of gravitational waves in our new model  $\beta$  we evaluate the volumetric merger rate at time  $t_j$  by simply summing over the metallicity bins and birth time bins,

$$R_{\text{GW}}^{\beta}(t_j) = \frac{1}{V_{\text{sim}}} \sum_{i=0}^{170} \sum_{k=0}^{100} \frac{d^3 N_{\text{GW}}^{\beta}}{dZ_k dt_i dt_j}, \quad (17)$$

where  $V_{\text{sim}}$  is the volume of our original SHARK simulation.

For efficiency, we choose a sample  $\tilde{N}_{\text{GW}}^{\beta} = 10^6$  binaries from our  $\beta$  COMPAS run to estimate the effective factor while using the full sample for our fiducial model,  $\alpha$ , where  $N_{\text{GW}}^{\alpha} = 170 \times 10^6$ .<sup>6</sup> The merger rate is reduced by a factor of 170 but this is compensated for by the new normalisation in Eq. 16,  $\frac{M_{\text{tot}}^{\alpha}}{M_{\text{tot}}^{\beta}} \sim 170$ . Thus, this formalism removes the dependence on the sample size. We test the validity of our rescaling procedure and using only  $10^6$  binaries for model  $\beta$  by splitting one of our fiducial COMPAS runs from Rauf et al. (2023) (using the remnant mass model of Mandel & Müller 2020) into 170 by  $10^6$  sub-samples. We then use these to estimate

<sup>6</sup> Note that we draw only  $\sim 10^6$  unique masses for this process from the IMF, but each one is randomly assigned a different metallicity from SHARK, resulting in  $\sim 170 \times 10^6$  unique combinations of initial masses and metallicities, and so  $\sim 170 \times 10^6$  unique binaries.



**Figure 2.** Volumetric BBH merger rate with remnant mass prescription from Mandel & Müller (2020) as a function of redshift. The blue line is the Rauf et al. (2023) merger rate from running COMPAS and SHARK. The red dashed line is the merger rate rescaled with the effective factor calculated using the full COMPAS run containing  $170 \times 10^6$  binaries. The grey lines (whose overlapping nature visually appears as a shaded black region) are the rescaled merger rate where the effective factor has been calculated using 170 unique sub-samples of  $10^6$  binaries from the full COMPAS run. The green line is the average over these sub-samples.

the rescaled merger rate for this model given  $\alpha = \text{Fryer et al. (2012)}$ , and look at the spread in the merger rate. We show in Fig. 2 that overall the rescaling method using equal numbers of binaries for models  $\alpha$  and  $\beta$  recovers the Rauf et al. (2023) merger rate from our high-fidelity cross-matching, but also that using only  $10^6$  binaries for model  $\beta$  returns, on average, an unbiased estimate of the merger rate. The use of a smaller number of binaries in the rescaling procedure does introduce additional Poisson scatter in the merger rate, which we propagate through into our model comparison/evaluation (see Section 4).

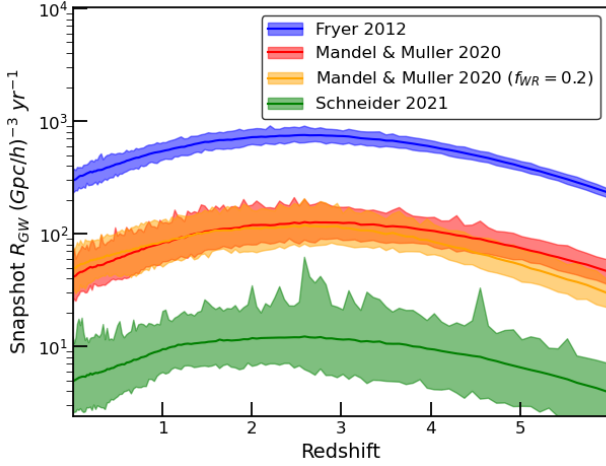
In Fig. 3, we show the volumetric BBH merger rates in the SHARK simulations using the effective factor approximation for four different models. In all cases we find that we are able to recover rates from Rauf et al. (2023) very well. We note that the plot is noisier than in Rauf et al. (2023) due to the smaller sample of BBHs taken from each COMPAS run, but again this noise arises only from Poisson sampling error.

## 4 BAYESIAN INFERENCE AND MODEL COMPARISON

In this section we outline our procedure for comparing different COMPAS models to data from GWTC-1 (Abbott et al. 2019a), GWTC-2 (Abbott et al. 2021a) and GWTC-3 (Abbott et al. 2023b) including our rescaling methodology. We validate this method by comparing the likelihoods given the data for the four models from Rauf et al. (2023) for which we have both ‘brute-force’ measurements of the GW rate and our rescaled calculations.

### 4.1 Selection criteria

The Bayesian analysis will depend strongly on  $N_{\text{obs}}$ , which is the number of confident GW detections. How confident we are in the



**Figure 3.** Volumetric rate plot as a function of redshift from Rauf et al. (2023) recreated with the effective factor scaling method. The error on the plots are the range of merger rates from sampling from each COMPAS run 170 times. We use a logarithmic scale for the y-axis, as the merger rates for the models spans several orders of magnitude.

GW signals and the validity of the sources is determined by the following ranking statistics:

(i) Signal-to-noise ratio (SNR): This is proportional to the signal amplitude divided by the noise amplitude. The higher the SNR, the stronger the signal compared to the noise, and more likely that the signal will be detected. This is the main ranking statistic when the signal shape is known from theory, such as the signal for a BBH merger. We use the network SNR, which is the square root of the sum of the squares of SNRs from the triggers which form the GW event (only from the detectors that produce triggers contributing to the event; Cabourn Davies & Harry 2022). The SNR,  $\rho$ , in a single detector is defined as,

$$\rho^2 = \langle h|h \rangle = 4\text{Re} \left[ \int_{f_{\min}}^{f_{\max}} \frac{h^*(f)h(f)}{S(f)} df \right], \quad (18)$$

where  $h(f)$  is the waveform measured by the detector,  $h^*(f)$  is its Hermitian conjugate,  $S(f)$  is the power spectrum density,  $\text{Re}[\ ]$  is the real component of the inner product, and  $f_{\min}$  and  $f_{\max}$  are the minimum and maximum limits for the considered frequency range (Farr et al. 2015; Barrett et al. 2018).

(ii) False alarm rate (FAR): Once the threshold for the SNR is set, the FAR can be determined, providing an alternative way to estimate the significance of the event (Zheng et al. 2021). The FAR is the number of noise events with a signal equal to or greater than the GW event, divided by the total duration of the background data. The false alarm rate of single-detector candidates given  $\rho$  is,

$$\text{FAR}(\rho) = \int_{\rho}^{\rho_{\max}} \Lambda_n p_n(\rho') d\rho', \quad (19)$$

where  $\Lambda_n$  is the mean Poisson rate of noise triggers and  $p_n(\rho')$  is the probability densities describing the distribution of  $\rho$  (Callister et al. 2017).

(iii)  $p_{\text{astro}}$ : The probability that a particular event is of astrophysical origin, as opposed to noise fluctuations or terrestrial, is defined by  $p_{\text{astro}}$ . It is measured by comparing the GW event rate to the background event rate for a fixed SNR and FAR. The probability of

a trigger,  $x$ , being astrophysical is,

$$p_{\text{astro}}(x) = \frac{p(x|\mathcal{S})\pi(\mathcal{S})}{p(x|\mathcal{S})\pi(\mathcal{S}) + p(x|\mathcal{N})\pi(\mathcal{N})}, \quad (20)$$

where  $\mathcal{S}$  and  $\mathcal{N}$  describe the signal and noise model, as well as the distribution of signal and noise sources respectively (Callister et al. 2017; Broekgaarden et al. 2023).

For our most rigorous selection, we combine  $\text{FAR} \leq 0.25 \text{ yr}^{-1}$  and  $p_{\text{astro}} \geq 0.9$ , as the FAR alone will not provide information on the sources being astrophysical, but rather the behaviour of the signal background. The  $p_{\text{astro}}$  reveals more events in the population rich areas of the parameter space, and is a more realistic metric for setting threshold for electromagnetic follow-up (Lynch et al. 2018; Andres et al. 2022). The caveat of  $p_{\text{astro}}$  is that independent pipelines will obtain different  $p_{\text{astro}}$  due to a difference in data reduction techniques. Banagiri et al. (2023) proposes a unified  $p_{\text{astro}}$  by combining triggers from all pipelines, which incorporates extra information about a signal. Galaudage et al. (2020) also emphasises the need for a pipeline independent  $p_{\text{astro}}$ , as they show that the posterior distribution of the population hyper-parameters can change with an astrophysically motivated prior. Given these caveats, we also investigate a simpler, broader selection using only an  $\text{SNR} \geq 8$  cut, which matches that used in evaluating our detection probability function  $p_{\text{det}}(\theta)$  for our models and provides us with a larger samples of events. However, the downside of this is that it potentially allows events into our sample that may not be astrophysical. In these cases, we allow our Bayesian framework to decide if these events should be ruled out.

Our supersample of GW events for querying the probability distribution function (PDF) for a set of parameters in each COMPAS model (see Section 4.2) consists of confirmed BBHs and BH-mass-gap events such as GW190521 (Abbott et al. 2020a) and GW190814 (Abbott et al. 2020c). This gives a total of 93 catalog GW detections. Of these, the following GW events are *not* included in our analysis for the following reasons:

- GW170817: Confirmed BNS (Abbott et al. 2017)
- GW190425: Confirmed BNS (Abbott et al. 2020b)
- GW190426\_152155: Not astrophysical (Abbott et al. 2021a)
- GW190909\_114149: Not astrophysical (Abbott et al. 2021a)
- GW191219\_163120: Confirmed NSBH (Abbott et al. 2023b)
- GW200115\_042309: Confirmed NSBH (Abbott et al. 2021b)

We then apply our two sets of selection criteria. We have 79 GW events with  $\text{SNR} \geq 8^7$  and 60 GW events with  $p_{\text{astro}} \geq 0.9$  and  $\text{FAR} \leq 0.25 \text{ yr}^{-1}$ . Table B1 lists the events used with each selection criteria. Out of the 87 BBH GW events we could have included (93 less the 6 listed above) the following GW events do not fit either criteria:

- GW190403\_051519
- GW190514\_065416
- GW191113\_071753
- GW200208\_222617
- GW200220\_061928
- GW200306\_093714
- GW200308\_173609
- GW200322\_091133

## 4.2 Fitting LVK posterior for various COMPAS models

Our basic fitting methodology is based on the widely-used Bayesian framework (e.g., Stevenson et al. 2015; Bouffanais et al. 2019; Mandel et al. 2019; Vitale et al. 2022; PÉrigois et al. 2023). Given the

<sup>7</sup> We use the network SNR provided in <https://gwosc.org/eventapi/html/GWTC/?page=1&pagesize=100>.

BBH GW data from some observing runs,  $\mathcal{H}$ , the posterior for a model specified via hyper-parameters  $\lambda$  (for instance a certain mass transfer efficiency  $f_{\text{MT}}$  or specific AM loss  $f_\gamma$ ), takes the form of an inhomogeneous Poisson distribution,

$$\mathcal{L}(\lambda, N_\lambda | \mathcal{H}) = \pi(\lambda) e^{-\mu_\lambda} N_\lambda^{N_{\text{obs}}} \prod_k \mathcal{I}^k. \quad (21)$$

$N_\lambda$  is the number of GW events predicted by the model in a given time-frame, while  $\mu_\lambda$  is the number the model suggests we would actually detect.  $N_{\text{obs}}$  is the actual number of detected events in our dataset. We assume that *a priori* all hyper-parameter combinations are equally likely, so  $\pi(\lambda) = 1$ . Finally,  $\mathcal{I}^k$  denotes the likelihood of observing GW event  $h^k$  given the hyper-parameters  $\lambda$ , written as a marginalisation over the binary properties predicted by that model,

$$\mathcal{I}^k = \int \mathcal{L}(h^k | \theta) p_{\text{sim}}(\theta | \lambda) d\theta = \frac{1}{N_s^k} \sum_{i=1}^{N_s^k} \frac{p_{\text{sim}}(\theta_i^k | \lambda)}{\pi^k(\theta_i^k)}, \quad (22)$$

where  $\theta = \{\mathcal{M}_c, q, z\}$  are the GW event parameters chirp mass, mass ratio and redshift respectively and  $p_{\text{sim}}(\theta | \lambda)$  is the probability distribution function for the event-level parameters generated by COMPAS using the SHARK metallicity as input. The latter sum arises from the use of the  $N_s^k$  publicly available posterior samples to represent the likelihood of event  $k$ ,  $\mathcal{L}(h^k | \theta)$ . Formally  $p_{\text{sim}}(\theta | \lambda)$  is obtained by marginalising over all the metallicities and birth times used that can give rise to an event with parameters  $\theta$ :

$$p_{\text{sim}}(\theta | \lambda) = \int_0^\infty \int_0^{t_j} p_{\text{sim}}(\mathcal{M}_c, q, z(t_j) | Z, t_i, \lambda) dt_i dZ. \quad (23)$$

$N_\lambda$  is the number of GW events in the Universe within an observing time that is expected from each COMPAS model. We can evaluate this given our rescaled merger rate density (expressed in individual metallicity and birth time bins, and in units of  $\text{Mpc}^{-3} \text{yr}^{-1}$ ) and  $p_{\text{sim}}$ ,

$$N_\lambda = 4\pi T_{\text{obs}} \int_{\theta \in V_\theta} \int_0^\infty \int_0^{t_j(z)} p_{\text{sim}}(\theta | Z, t_i, \lambda) R_{\text{GW}}^\lambda(Z, t_i, t_j(z)) \chi^2(z) \frac{c}{H(z)} \frac{1}{1+z} dt_i dZ d\theta, \quad (24)$$

where  $\chi(z)$  is the comoving distance to the redshift  $z$ ,  $c$  is the speed of light,  $H(z)$  is the Hubble parameter, and  $T_{\text{obs}}$  is the total LVK observing time (duty cycle). This is 118.0, 138.8 and 142.0 days for O1, O2 and O3 respectively, which we converted to years.

We can extend this to the predicted *detected* number of GW events using a detection probability for the event-level parameters  $p_{\text{det}}(\mathcal{M}_c, q, z)$  given a particular detector configuration and signal-to-noise threshold.

$$\mu_\lambda = 4\pi T_{\text{obs}} \int_{\theta \in V_\theta} \int_0^\infty \int_0^{t_j(z)} p_{\text{sim}}(\theta | Z, t_i, \lambda) p_{\text{det}}(\theta) R_{\text{GW}}^\lambda(Z, t_i, t_j(z)) \chi^2(z) \frac{c}{H(z)} \frac{1}{1+z} dt_i dZ d\theta. \quad (25)$$

For Eq. 21, we are assuming that  $N_\lambda$  events are all formed via the isolated formation channel. There is some evidence for some events in GWTC-3 (Abbott et al. 2023b) forming via the dynamical channel (Arca Sedda 2021; Romero-Shaw et al. 2022). To avoid any assumptions on the formation channel and discount information on the predicted vs. observed GW rate, we can marginalise over  $N_\lambda$  to remove its dependence by multiplying with a prior  $\pi(N_\lambda) = \frac{1}{N_\lambda}$  and

integrating (Fishbach et al. 2018).

$$\begin{aligned} \mathcal{L}(\lambda | \mathcal{H}) &= \int \frac{1}{N_\lambda} e^{-N_\lambda \beta(\lambda)} N_\lambda^{N_{\text{obs}}} \prod_k \mathcal{I}^k dN_\lambda \\ &= (N_{\text{obs}} - 1)! \prod_k \frac{\mathcal{I}^k}{\beta(\lambda)}, \end{aligned} \quad (26)$$

For convenience we can use  $N_\lambda$  and  $\mu_\lambda$  to define some detection efficiency  $\beta(\lambda) = \frac{\mu_\lambda}{N_\lambda}$ . We make model inferences using both of these posteriors in our work and compare how our conclusions change when  $N_\lambda$  is included as information.

In order to evaluate our posteriors, we first calculate the model volumetric rate using a variant of Eq. 17 but kept in separate metallicity and birth time bins. Given  $T_{\text{obs}}$ , the merger rate and the appropriate volumetric factors, we then find the number of GW events that would fall in each metallicity and time bin,  $N_{\lambda,i,j}$ , and draw that many events (characterised by a chirp mass and mass ratio) from our full list of  $10^6$  COMPAS binaries.<sup>8</sup> We store the chirp mass, mass ratio and merger redshift of the events across all metallicity and birth time bins before then building the PDF  $p_{\text{SIM}}(\theta | \lambda)$  using Kernel Density Estimation (KDE), in particular using the fast KDEpy.FFTKDE package. The bandwidth for our KDE is calculated using Silverman's rule of thumb (Silverman 1986) for each hyper-parameter distribution and averaged to obtain the single band-width for the full 3D distribution. As a by-product of method we can also estimate  $N_\lambda$  and  $\mu_\lambda$  at the same time by simply keeping track of how many COMPAS binaries we are drawing.

For  $p(\theta_i^k | \lambda)$ , we query the KDE-built PDF using the chirp mass, mass ratio and redshift from the public LVK posterior samples. We construct  $\pi^k(\theta_i^k)$  by building another KDE with the public LVK prior samples and querying this again with the posterior samples at each point  $\theta_i^k$ . In the case where  $p_{\text{sim}}(\theta | \lambda) = 0$  for all posterior samples, this implies that the COMPAS model cannot reproduce the BBHs in the posterior samples, as they fall outside of the KDE. We flag these events and ultimately remove them from the analysis for that specific model comparison. For each of the  $N_{\lambda,i,j}$ , we calculate the detection probability,  $p_{\text{det}}$ , based on the masses, redshift and luminosity distance. The detection probability comes from a selection effect function in COMPAS, which takes the BH masses, redshift and luminosity distance as inputs (Barrett et al. 2018; Team COMPAS: Riley, J. et al. 2022). The SNR is calculated by computing the source waveform using the waveform model from the LAL software suite (LIGO Scientific Collaboration et al. 2018).

We apply the same procedure also to estimate marginalised, observed 1D distributions  $p_{\text{SIM\_DET}}(\mathcal{M}_c | \lambda)$ ,  $p_{\text{SIM\_DET}}(q | \lambda)$  and  $p_{\text{SIM\_DET}}(z | \lambda)$ .<sup>9</sup> We show these 1D distributions for our four fiducial models in Fig. 4 alongside the GWTC data, with our main observations being:

- With the exception of Schneider et al. (2021), the models replicate the first peak of the observed chirp mass distribution well. No models fit the second peak around  $30M_\odot$ . For all models, there is zero probability of producing chirp masses  $> 40M_\odot$  until  $100\text{--}130M_\odot$ . This implies these models predict high mass BHs that either may

<sup>8</sup> We find  $N_{\text{COMPAS},i,j}$  in each metallicity and snapshot bins from the COMPAS run. We calculate a weight defined as  $N_{\lambda,i,j}/N_{\text{COMPAS},i,j}$  to avoid repeatedly drawing from the COMPAS run. We apply this weight to the PDF. When calculating the detected PDF, we multiply this weight by  $p_{\text{det}}$ .

<sup>9</sup>  $p_{\text{SIM\_DET}}$  refers to the detected PDF, which is marginalised over the other two fitting parameters *after* the detection probability is applied.

have not been detected due to the rarity of finding these types of events with current detectors, as they may be associated with noise in the low frequency band, or as a result of inaccuracies in the models resulting in an incorrect representation of the observed population.

- The mass ratio distribution is quite noisy but overall relatively flat, similar to the observed distribution. However for [Mandel & Müller \(2020\)](#) and [Schneider et al. \(2021\)](#) there is a peak in the distribution after  $q \approx 0.8$  indicating a preference for equal mass binaries. [Mandel & Müller \(2020\)](#) with  $f_{\text{WR}} = 0.2$  has a preference for  $q \sim 0.6$  and [Fryer et al. \(2012\)](#) for  $q \sim 0.7 - 0.8$ .

- The shape of the redshift distributions for the models fit the observed distribution quite well. Overall, the [Fryer et al. \(2012\)](#) model is the best fit, while the other models have a wider distribution. This indicates BBHs from these models are more likely to be detected at higher redshifts (closer to  $z \sim 1$ ) compared to [Fryer et al. \(2012\)](#).

### 4.3 Quantifying the uncertainty in $N_\lambda$ , $p(\lambda, N_\lambda|\mathcal{H})$ and $p(\lambda|\mathcal{H})$

For our four models in [Rauf et al. \(2023\)](#) we were able to split COMPAS outputs into 170 sub-samples to compute 170 effective factors and produce 170 theoretical,  $N_\lambda$ , and detected,  $\mu_\lambda$ , numbers of GW events. We show the distribution of the detected number of events in Fig. 5. For the error on the total number of events, we can use the standard deviation across our 170 subsamples. However, if we have only one sample of  $N_{\text{sample}}$  binaries for a COMPAS output, we require an alternative approach to assign an error for these models. Given the number of models we test in this work, the alternatives of simply using more samples or bootstrapping using the  $N_{\text{sample}} = 10^6$  samples would be computationally challenging. Furthermore, sub-sampling or bootstrapping to generate an error would fail to capture the impact of progenitors of rare events that are not represented in the  $N_{\text{sample}} = 10^6$ , but which are present in the distributions obtained from our four models with  $N_{\text{sample}} = 170 \times 10^6$ , and from which our errors are extrapolated. We find that the ratio of the standard deviation to the mean number of events as a function of  $\log N_\lambda$  for the four models has a linear trend such that the relative error increases as the total number of GW events generated by the model decreases, in a roughly Poissonian fashion. In Fig. 6 we show the linear fits for the standard deviation in  $N_\lambda$  and  $\mu_\lambda$ , which we show in the following equations:

$$\frac{\Delta N_\lambda}{N_\lambda} \approx -0.04 \log N_\lambda + 0.55 \quad (27)$$

$$\frac{\Delta \mu_\lambda}{\mu_\lambda} \approx -0.03 \log N_\lambda + 0.43 \quad (28)$$

We can now apply this to single samples, assuming similar Gaussian distributions for other COMPAS models. To justify this, we find that the mean  $\mu_\lambda$  over the 170 sub-samples in Fig. 5 is equivalent to the  $\mu_\lambda$  we obtain using the [Rauf et al. \(2023\)](#) merger rate from the full COMPAS output (one sample) for all four models that we test.

Now to address the uncertainty in  $\mathcal{I}^k$ ; We find another linear trend for the standard deviation in  $\sum_k^{N_{\text{obs}}} \log \mathcal{I}^k$  over the 170 sub-samples as a function of the mean  $\log N_\lambda$  for both our selection criteria, as shown in Fig. 7.

The linear fits for the standard deviation in the likelihood sums are given by:

$$\frac{1}{N_{\text{obs}}} \Delta \sum_k^{N_{\text{obs}}} \log \mathcal{I}^k \approx -0.05 \log N_\lambda + 0.70 \quad (29)$$

For new COMPAS models shown in Section 5, we are able to use

Eq. 29 and to calculate the error on  $\sum_k^{N_{\text{obs}}} \log \mathcal{I}^k$  given  $N_\lambda$  without needing to resample. We can now propagate the error in  $N_\lambda$ ,  $\mu_\lambda$  and  $\sum_k^{N_{\text{obs}}} \log \mathcal{I}^k$  into the error on the log-likelihood.

$$\begin{aligned} \Delta \log \mathcal{L}(\lambda|\mathcal{H}) &= \sqrt{\left(N_{\text{obs}} \frac{\Delta \mu_\lambda}{\mu_\lambda}\right)^2 + \left(N_{\text{obs}} \frac{\Delta N_\lambda}{N_\lambda}\right)^2 + \left(\Delta \sum_k^{N_{\text{obs}}} \log \mathcal{I}^k\right)^2} \\ \Delta \log \mathcal{L}(\lambda, N_\lambda|\mathcal{H}) &= \sqrt{(\Delta \mu_\lambda)^2 + \left(N_{\text{obs}} \frac{\Delta N_\lambda}{N_\lambda}\right)^2 + \left(\Delta \sum_k^{N_{\text{obs}}} \log \mathcal{I}^k\right)^2} \end{aligned} \quad (30)$$

Although only rudimentary given the small number of models available to us to generate fitting formulae, even adopting this simple uncertainty on the likelihood allows us to consider the statistical variations in model evaluation using population synthesis. As such it enables us to draw more confident conclusions on which hyper-parameters cause major changes in the likelihood, and whether these changes are physically significant.

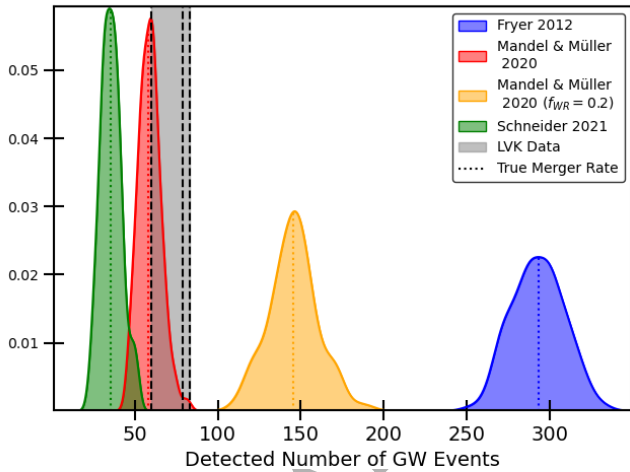
### 4.4 Comparing posteriors with and without rescaling merger rate

Fig. 8 shows the  $\mathcal{I}^k$  term for individual GW events. This is a practical approach to determine the probability of individual GW events occurring in the populations generated by the COMPAS models. The [Schneider et al. \(2021\)](#) model generally returns high probabilities when querying the PDF, while the [Fryer et al. \(2012\)](#) model is disfavoured. We find some events are outliers and are not well reproduced by some or all of the models relative to the other events. As an example, [Mandel & Müller \(2020\)](#) with  $f_{\text{WR}} = 0.2$  and [Schneider et al. \(2021\)](#) are less likely to output the progenitors of GW190814, GW190917\_114630 and GW200210\_092254 compared to [Fryer et al. \(2012\)](#) and [Mandel & Müller \(2020\)](#). For GW190426\_190642, GW190602\_175927 and GW191109\_010717 [Schneider et al. \(2021\)](#) is distinctly preferred over the other models, although in these cases, and especially that of GW190521, the models are generally less able to produce the required event-level properties. This is expected for GW190521 as both masses,  $\sim 85M_\odot$  and  $\sim 66M_\odot$ , most likely lie in the pair-instability supernovae mass gap, making it atypical of binaries formed via the isolated formation channel ([Abbott et al. 2020a](#)).

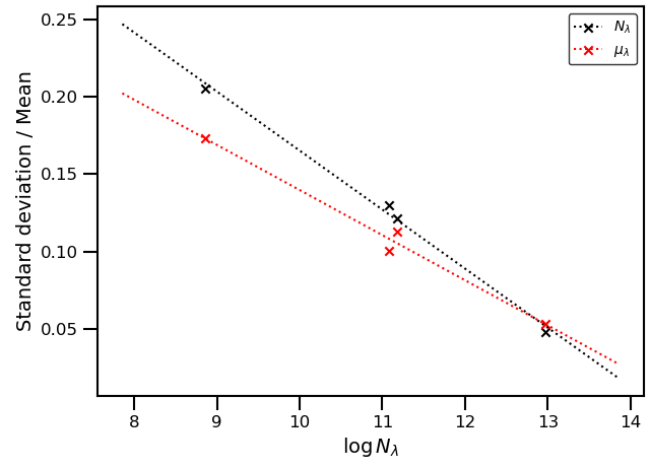
For a more comprehensive assessment of the models, we can analyse the likelihoods. In Fig. 9 we now incorporate the detection probability and show the likelihoods with and without  $N_\lambda$  dependence. The [Mandel & Müller \(2020\)](#) appears to the best fit to the observed population. This can be justified with Fig. 4. The [Mandel & Müller \(2020\)](#) model has two peaks in the chirp mass distribution and a relatively flat mass ratio distribution, similar to the LVK population distribution. If we assume that the models output accurate populations, then incorporating the number of GW events makes the likelihood worse – [Fryer et al. \(2012\)](#) greatly overestimates the number of detected events, while [Schneider et al. \(2021\)](#) underestimates it. In general the latter is preferred as it leaves ‘room’ for some of the LVK observed events to arise from the dynamical formation channel. The exception is the [Mandel & Müller \(2020\)](#) model, which performs well since the model predicts the same number of GW events within the observed range. The true number of events predicted is  $\sim 60$ , comparable to  $N_{\text{obs}}$  for the  $p_{\text{astro}} \geq 0.9$  and  $\text{FAR} \leq 0.25\text{yr}^{-1}$  selection criteria. For  $\text{SNR} \geq 8$ , we see an increase in the likelihood for all models, but the distributions remain the same. The biggest difference in the unmarginalised likelihood for our two selections occurs for the [Fryer](#)



**Figure 4.** Distribution of chirp mass, mass ratio and redshift for different models. Different colours refer to the different remnant mass prescriptions investigated in Rauf et al. (2023). The grey shaded regions are the inferred distribution from the LVK collaborations, calculated using the posterior samples for the  $N_{\text{obs}} = 79$  GW events based on the  $\text{SNR} \geq 8$  cut.



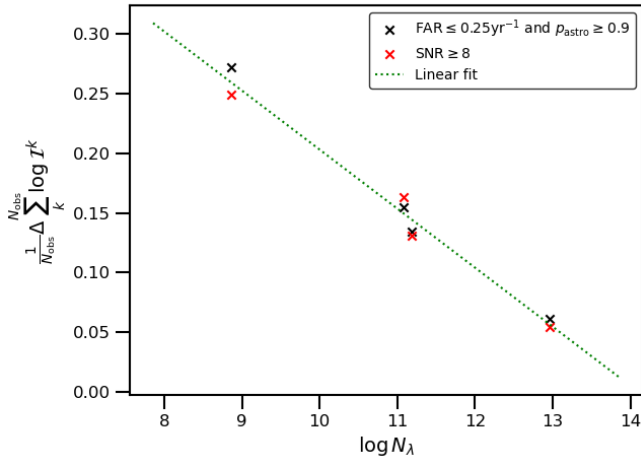
**Figure 5.** Normalised histograms of the detected number of BBH merger events calculated using Eq. 25 over the LVK observing time for the last three observing runs. The coloured distributions are obtained from the predictions from the model for 170 sub-samples. The grey band represents the uncertainty in the observed number of BBH mergers depending on the selection criteria adopted (see Section 4.1). The dotted lines are the number of events predicted using the Rauf et al. (2023) merger rate from Rauf et al. (2023).



**Figure 6.** Relative error on the numbers of generated and detected events versus the average  $\log N_\lambda$  for the four Rauf et al. (2023) models from 170 sub-samples. The crosses correspond to the relative error from the models and the dashed lines are linear fits to the points. The black points and fit are for  $N_\lambda$  and the red points and fits are for  $\mu_\lambda$ .

et al. (2012) model, which produces the largest number of detected events out of the four models. This is due to  $N_{\text{obs}}$  increasing to 79,

such that the difference between the predicted  $\mu_\lambda$  and the number of observed LVK events is slightly reduced.



**Figure 7.** The standard deviation in the sum of the log-likelihood per event,  $\log \mathcal{I}/N_{\text{obs}}$ , versus the mean  $\log N_\lambda$  over the 170 sub-samples for the four Rauf et al. (2023) models. We show the linear fit (green dashed line) that applies to both sets of selection criteria; FAR  $\leq 0.25\text{yr}^{-1}$  and  $p_{\text{astro}} \geq 0.9$  (black) and SNR  $\geq 8$  (red).

## 5 RESULTS

Having validated our methodology using the four models analysed in Rauf et al. (2023), this section discusses the outputs from our new set of COMPAS models where we vary the range of hyper-parameters described previously in Section 2 (see also Table 1). We start with describing the trends in the number of detected events when varying the COMPAS hyper-parameters. We calculate the Bayes factor to identify groups of models that are more preferred. We then show our likelihood analysis, where we discuss the best and worst models when compared to the observed population parameters. We follow up this analysis and look at specific events and their  $\mathcal{I}^k$  value to determine which models are more and less likely to generate the progenitors of these events. See Tables C1, C2 and C3 in Appendix C for a detailed description of the hyper-parameters for all 304 COMPAS models in this analysis.

### 5.1 $\mu_{\text{GW}}$ versus COMPAS parameters

Since we explore the number of GW events as a parameter in our posterior analysis, in this section we comment briefly on the trends between the detected number of GW events,  $\mu_\lambda$ , and the COMPAS hyper-parameters.

In the left plot of Fig. 10, we see  $\mu_{\text{GW}}$  increase with mass transfer fraction,  $f_{\text{MT}}$ , when the corresponding AM is lost directly from the accretor ( $f_\gamma = 0$ ). Increasing the mass transfer can lead to more equal mass objects, which we find most of our COMPAS models producing. This trend disappears for higher  $f_\gamma$ . We credit this to non-trivial evolutionary effects such as the stability prescriptions, that determine any future interactions within the binary. This preference for intermediate values of  $f_\gamma$  agrees with the findings by Santoliquido et al. (2021). Irrespective of  $f_{\text{MT}}$ ,  $\mu_{\text{GW}}$  peaks for moderate values of  $f_\gamma = 0.5$  but drops again for high  $f_\gamma$  when AM loss occurs closer to the L2 point. At this point, mass is lost at higher velocities so more AM is lost from the system, which in turn can induce CE evolution and merger during the mass transfer stage, stopping the formation of a BBH. These results are in agreement with Lu et al. (2023); van Son et al. (2022b) that the number of stable mass channel objects

increases for decreasing  $f_\gamma$ , as opposed to the objects formed by CE evolution. Marchant et al. (2021) also find that the stable channel dominates the BBH merger rate. We note that when  $f_{\text{MT}} = 1$ ,  $f_\gamma$  should have no impact. Any variations in the number of events at  $f_{\text{MT}} = 1$  is hence due to Eddington-limited accretion occurring after the first mass-transfer event(s), once the accretor has become a BH. <sup>10</sup>

We show the models that produce the closest  $\mu_\lambda$  to the observed LVK events in the right plot of Fig. 10. We find that for  $\mu_{\text{GW}}$  to be below 100 events (closer to the number of detected events from all observing runs), a CE efficiency  $\alpha_{\text{CE}} = 10, 100$  and  $f_\gamma = 0.75, 1.0$  with no CHE is required. This implies that the CE is ejected earlier in the evolution and the orbit tightens at a wider separation, which means longer merger times and less events in the same observed time. This result validates the studies by Santoliquido et al. (2021), Bouffanais et al. (2021) and Grichener (2023). As discussed,  $\alpha_{\text{CE}} = 100$  is extreme, which suggests an alternative approach for the CE is required for forming populations with a comparable number of GW events to LVK. With the exception of these models, the CE efficiency has no significant impact on the populations, as the range for the number of events does not vary significantly with  $\alpha_{\text{CE}}$ .

In Figure 11,  $\mu_{\text{GW}}$  decreases with increasing  $f_{\text{WR}}$ . We find this for any CHE mode and remnant mass prescription, but we show this for Schneider et al. (2021). This is due to increasing the mass loss rates leading to smaller BHs and wider separations, so less mergers occur in the same observing time. These predictions agree with the work by Riley et al. (2021) and Stevenson & Clarke (2022) using COMPAS. We find that for no CHE, the rate of change in  $\mu_\lambda$  across  $f_{\text{WR}}$  is smaller compared to the pessimistic and optimistic cases.

The pessimistic and optimistic CHE modes produced  $> 1000$  events due to their tight orbits. This suggests the CHE channel is not ideal for generating the current LVK detected population. We verify this with the likelihoods in Section 5.2.

### 5.2 Bayes Factor and Likelihood Analysis

To identify groups of models that are more preferred than others, we use the Bayes factor,  $\mathcal{B}$ . Since the priors for all models are the same, the Bayes factor is the ratio of the likelihoods between two models, here denoted  $\alpha$  and  $\kappa$ :

$$\mathcal{B}(\lambda|\mathcal{H}) = \frac{\mathcal{L}(\lambda_\alpha|\mathcal{H})}{\mathcal{L}(\lambda_\kappa|\mathcal{H})}; \quad \mathcal{B}(\lambda, N_\lambda|\mathcal{H}) = \frac{\mathcal{L}(\lambda_\alpha, N_\lambda|\mathcal{H})}{\mathcal{L}(\lambda_\kappa, N_\lambda|\mathcal{H})}. \quad (31)$$

In Fig. 12, we show the Bayes factor with the likelihoods using the GW events with the SNR  $\geq 8$  cut.

We note that the Bayes factor indicates there is strong evidence that some groups of models are preferred more than others, which we will further discuss in the likelihood analysis. We emphasise these with highlighted boxes in Fig. 12 labeled A-G. We can understand the range of Bayes factor by considering the different components that enter Eq. 31. The difference in  $\log \beta(\lambda)$  between different models ranges from  $\sim 0.05 - 3.44$ . However, we find that the  $\mathcal{I}^k$  terms are distributed over 50 orders of magnitude (see bottom plot in Fig. 18) and therefore dominates the Bayes factor. This confirms that some of the models are distinguishable and strongly preferred (more likely

<sup>10</sup> As a result of this finding, dual functionality in the AM loss was later implemented in COMPAS, as discussed in Section 2.4. If specific AM loss is set at the compact object as well as prior to the formation of the first compact object we expect the number of events will no longer vary for  $f_{\text{MT}} = 1$ , although at time of writing this functionality is yet to be comprehensively tested.

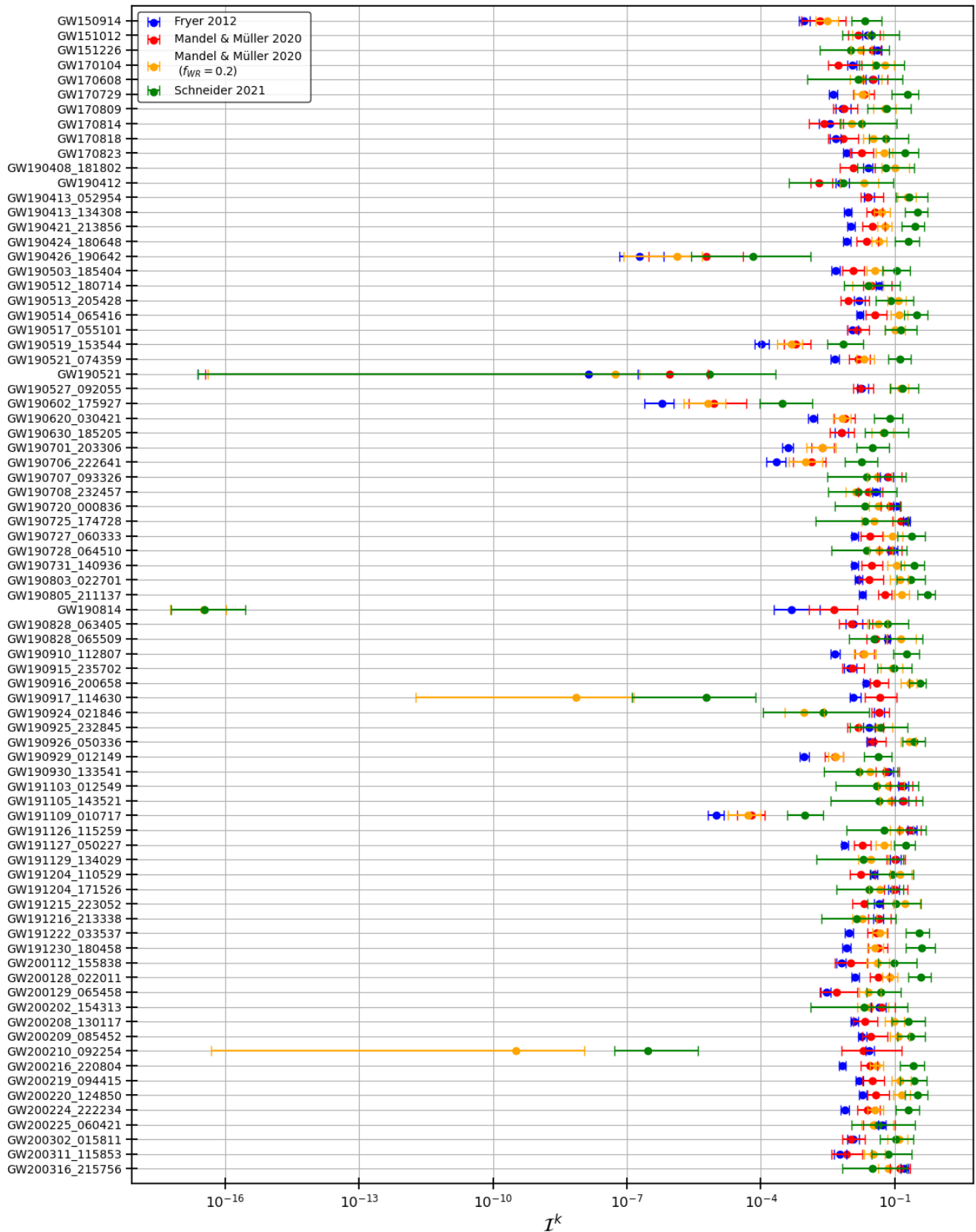
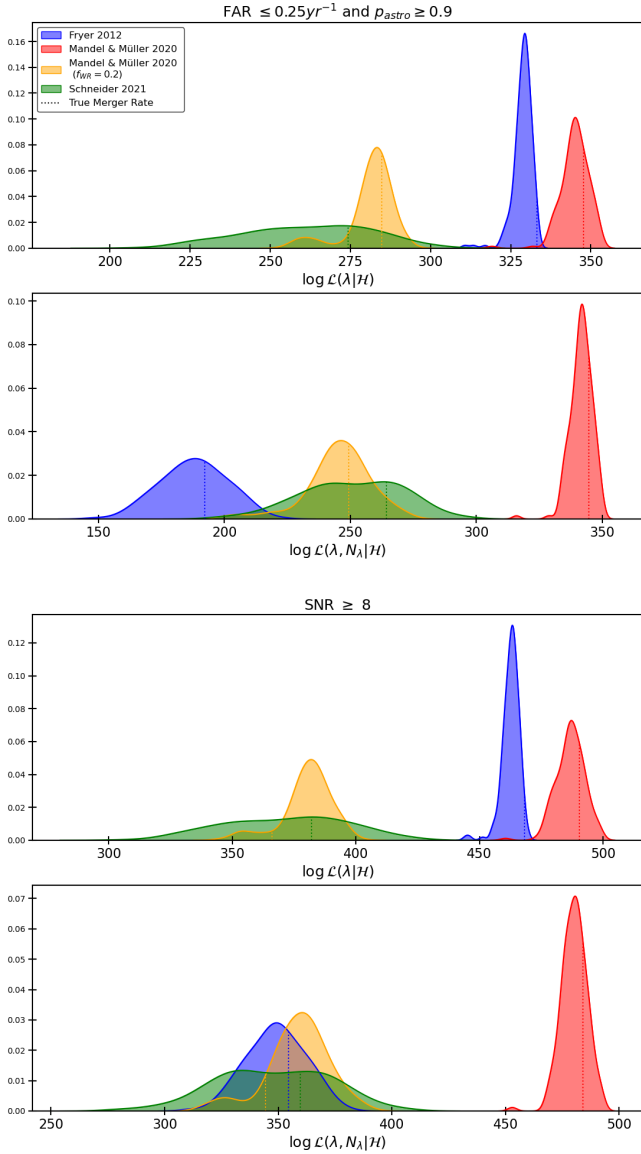


Figure 8. The likelihood,  $\mathcal{I}^k$ , for each GW event with SNR  $\geq 8$ . Each colour represents a different remnant mass prescription from COMPAS and the error bar represents the sampling error over 170 samples when building the simulated PDF of  $\mathcal{M}_c$ ,  $q$  and  $z$  given the model.



**Figure 9.** Normalised histograms of the log likelihoods obtained using 170 sub-samples for the models explored in Rauf et al. (2023) with and without the merger rate dependence. The top two panels are the likelihoods comparing the models to the GW events that fit the  $\text{FAR} \leq 0.25\text{yr}^{-1}$  and  $p_{\text{astro}} \geq 0.9$  criteria, while the bottom two panels compares to the  $\text{SNR} \geq 8$  criteria. The dotted lines are the likelihoods using the Rauf et al. (2023) merger rate. In all cases we see that the true likelihood is well represented within our distribution of 170 subsamples, indicating our use of a relatively smaller number of  $10^6$  binaries is not expected to bias our results.

to produce progenitors of particular GW events) over others purely based on the masses and redshift distributions they predict. When including  $N_\lambda$  as a parameter in the likelihood, the  $-\mu_\lambda + N_{\text{obs}} \log(\mu_\lambda)$  terms this introduces into the log-likelihood has significant weight, and comes to dominate the Bayes factor, meaning that a comparison of two models using this likelihood strongly favours those that produce a number of events close to the observed number of events in our data sample, even if the event-level distributions are less well matched. Hence, we carefully consider the results using both of our likelihoods.

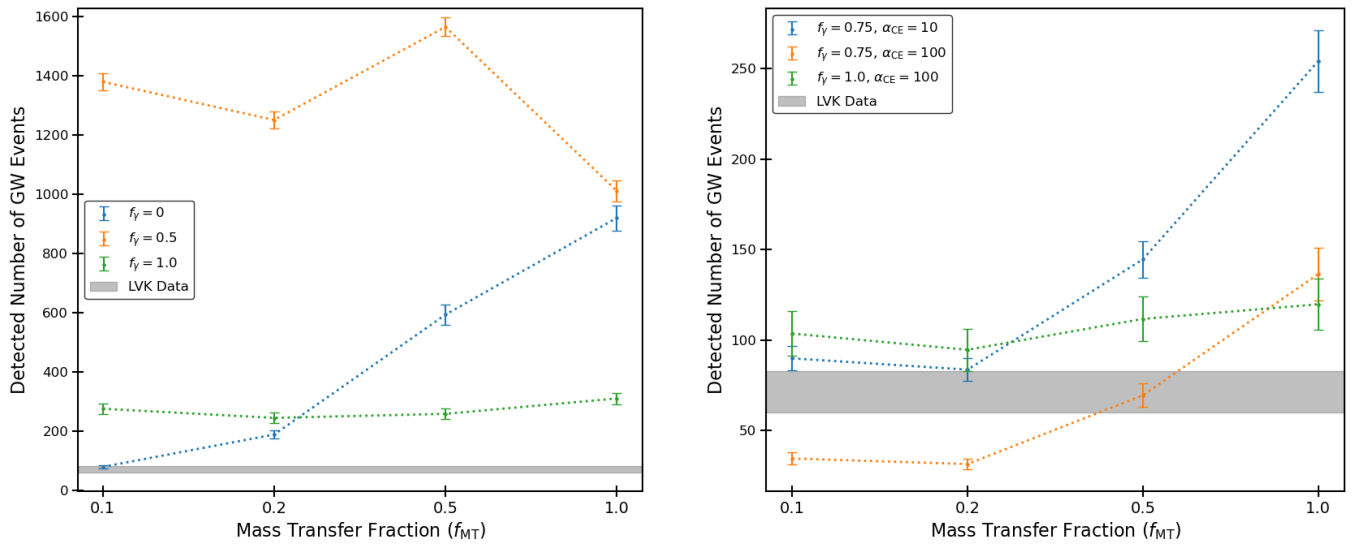
Having established groups of models that are strongly preferred over other models, we now discuss the likelihoods for individual cases and the trends within these groups. The following results (Fig. 13, 14, 15, 16, 17, 18) discussed use the  $\text{SNR} \geq 8$  selection cut. We see an increase in the likelihood, due to an increase in  $N_{\text{obs}}$ , across all models compared to using  $p_{\text{astro}} \geq 0.9$  and  $\text{FAR} \leq 0.25\text{yr}^{-1}$ , but the trends don't change substantially with the selection criteria and we find the differences are made more distinct when using the  $\text{SNR} \geq 8$  sample. We show this in Tables C1, C2 and C3 with Bayes factors between our various models.

We see an increase in the likelihood from  $f_\gamma = 0$  to  $f_\gamma = 0.5$  and we find that the likelihoods are similar with no CHE and pessimistic and optimistic CHE. In Fig. 13, we show the likelihoods for fixed  $f_\gamma$ . It is clear that  $f_\gamma = 0.5$  is preferred over  $f_\gamma = 1.0$ . Less BBH mergers are produced with  $f_\gamma = 1.0$ . Models with  $f_\gamma = 0.5$  produce more equal mass binaries, which is more consistent with the LVK population (Abbott et al. 2019b). However,  $f_\gamma = 0.75$  is unique in that for low mass transfer  $f_{\text{MT}} < 0.5$  the population is disfavoured similar to  $f_\gamma = 1.0$ , but for high mass transfer  $f_{\text{MT}} > 0.5$  the model is more favourable. In all cases, we see little impact from the choice of CE efficiency. This highlights the degeneracy between stellar modelling assumptions and populations. Ultimately, we find that in order to reproduce the observed events from LVK, we need either momentum loss to occur close to the accretor, or, if occurring further away, for the mass transfer fraction to be high. These trends are also highlighted by boxes A and B in Fig. 12, where moderate values of  $f_\gamma$  are strongly favoured when comparing population parameters.

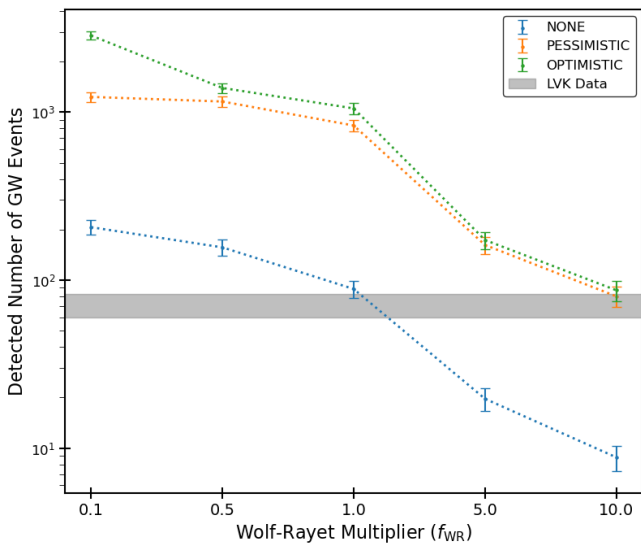
Fig. 14 then explores different combinations of the remnant mass prescription, CHE mode and Wolf-Rayet multiplier. Overall, we see weaker trends here than with the mass transfer fraction. However, the remnant mass prescription can be seen to impact the likelihood, with the Fryer et al. (2012) or Mandel & Müller (2020) prescriptions preferred. It is worth reiterating that, of these two, the Fryer et al. (2012) was seen to greatly over-predict the number of observed events to date. We find a weak trend of increasing likelihood in the absence of CHE, most evident for the Schneider et al. (2021) model, but overall less significant an impact than was seen for changes in the mass transfer and AM loss parameters. We also see a weak preference for larger  $f_{\text{WR}}$  depending on the remnant mass model.

### 5.2.1 Likelihood including $N_\lambda$

When including the dependence on the number of GW events, we are able to identify more distinct groups of trends, particularly with the CHE modes as shown in Fig. 15. This is in contrast to the likelihoods without the number of events, indicating that the CHE mode more strongly affects just the number of BBH mergers, rather than the underlying distributions of mass or redshift. The likelihood generally reduces when switching from no CHE to Pessimistic or Optimistic, which is not as notable for the likelihood without  $N_\lambda$  dependence. We also see this impact in Fig. 12, where box C and D correspond to no CHE being strongly preferred over Pessimistic and Optimistic CHE and box E implies Pessimistic is strongly preferred over Optimistic CHE. Allowing for CHE would mean tighter orbits and more mergers within Hubble time, which would over-predict the number of observed events when compared to  $N_{\text{obs}}$ . In addition, the likelihood now decreases from  $f_\gamma = 0.25$  to  $f_\gamma = 0.5$ . This is also due to the increased number of GW events, shifting them further from the observed number of LVK events. It then increases again with further increasing  $f_\gamma = 0.75, 1.0$  due to the reduced numbers of events. Within each choice of CHE, we see a preference for low mass transfer efficiency for  $f_\gamma = 0$ . This is required to keep the predicted number



**Figure 10.** Detected number of BBH merger events versus the mass transfer fraction for no CHE at fixed  $f_\gamma$ . *Left:* We fix the CE evolution parameter  $\alpha_{CE} = 0.1$  and show results for different specific AM loss parameters. *Right:* We show the models for any  $f_\gamma$  and  $\alpha_{CE}$  which output  $\mu_\lambda$  close to the number of observed LVK events given by the grey band.



**Figure 11.** Detected number of BBH merger events versus the Wolf-Rayet multiplier for fixed CHE modes and the Schneider et al. (2021) remnant mass prescription. The grey band is the observed number of LVK events.

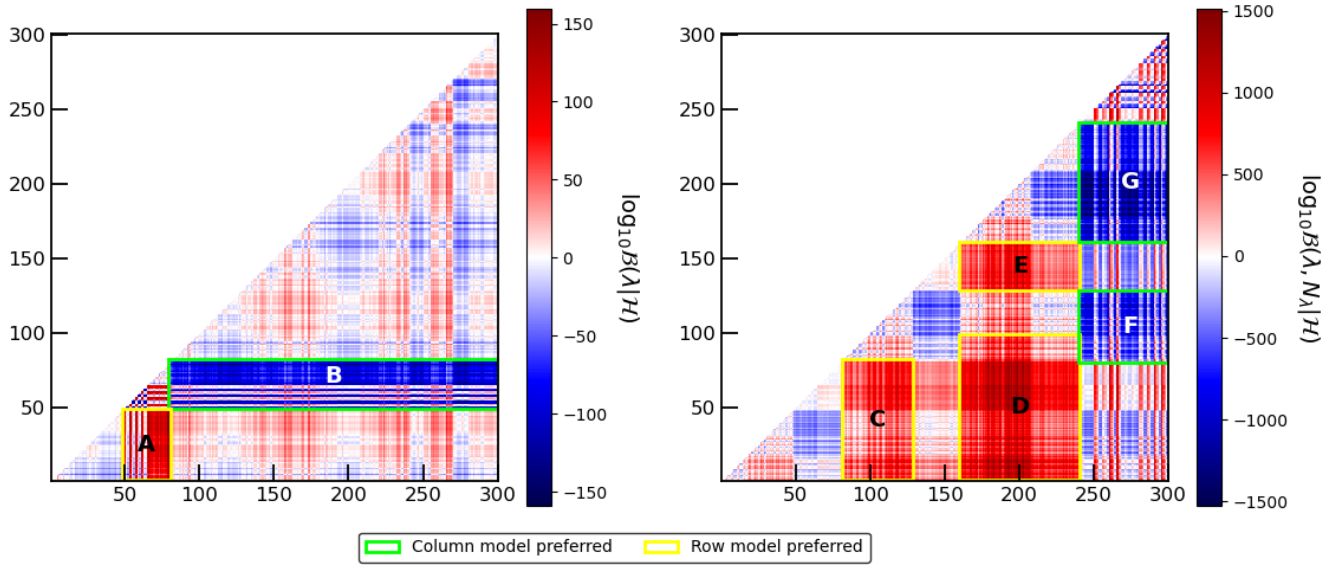
of events low or comparable to the actual number of LVK events, as these produce unequal mass wide binaries with long merger times.

Although no CHE is preferred, this can depend on  $f_\gamma$ . We show a subset of no CHE with varying  $f_\gamma$  in Fig. 15. Again we note the degeneracy in the hyper-parameters and populations. Low values of  $f_{MT}$  and  $f_\gamma = 0.25$  produce similar populations to  $f_\gamma = 0$ , while models with higher values of  $f_{MT}$  at  $f_\gamma = 0.25$  appear similar to Pessimistic CHE with  $f_\gamma = 0$  populations. Conversely, further increases to  $f_\gamma = 0.5$  and high  $f_{MT}$  produce high-likelihood populations similar to  $f_\gamma = 0$ , with now a decreasing likelihood as  $f_{MT}$  decreases.

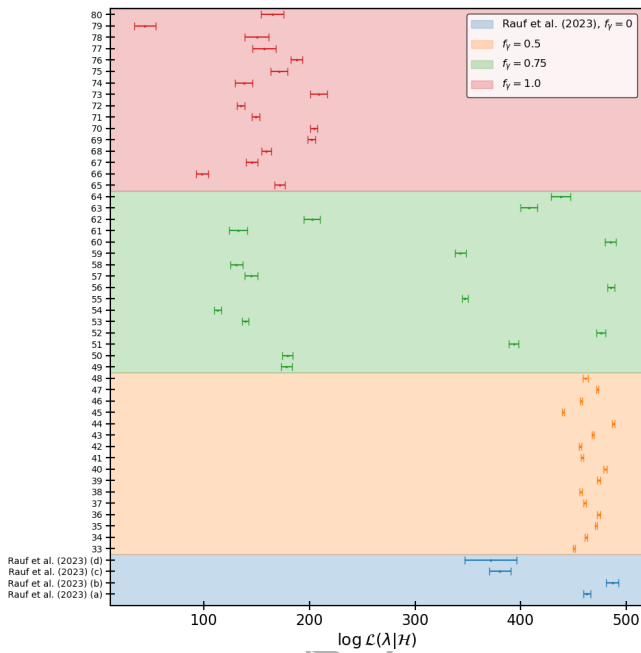
The likelihood further increases with  $f_\gamma = 0.75$  particularly again for high  $f_{MT}$ . There is hence a flip in the behaviour of increasing mass transfer fraction depending on whether the momentum loss is occurring closer to the accretor or L2 point. These changes in likelihood closely reflect the trends seen in the number of predicted events seen in Fig. 10, but further understanding of this complex behaviour warrants further study.

Finally, in Fig. 16, we find a strong case for likelihood increasing with  $f_{WR}$ , with the exception of Schneider et al. (2021) with no CHE. This is also shown in Fig. 12 by boxes F and G, where  $f_\gamma = 0$  and thermal timescale mass transfer rates are strongly preferred over moderate values of  $f_\gamma$  and fixed  $f_{MT}$ . For fixed CHE and remnant mass prescription, there is a preference for models with  $f_{WR} = 10$ . This is because mass loss via strong winds can widen binaries. This would increase the merger time and decrease the number of GW events closer to the LVK detected number of events. Overall, again no CHE is preferred over pessimistic and optimistic CHE — in the presence of CHE, one needs a rather finely-tuned set of: high Wolf-Rayet factor, high  $f_\gamma$  and low mass transfer efficiency to recover a reasonable number of mergers.

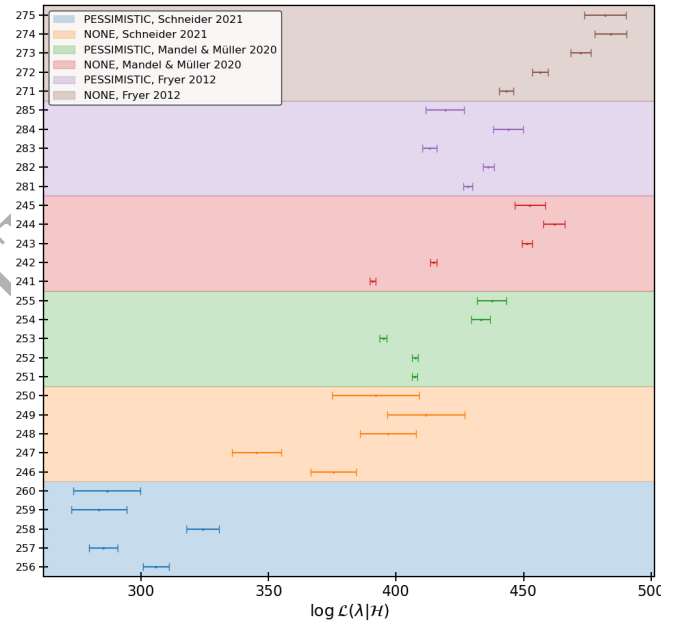
Given the main assumption in this paper being all GW events originate from the isolated binary formation channel, the observed number of events in the LVK data places an upper bound on  $\mu_\lambda$ . Any models that output  $\leq N_{obs}$  cannot be ruled out. However, even in the case where the number of detected events is close to  $N_{obs}$ , this is not a good indicator of if the model population will match the observed population. We propose checking the marginalised and unmarginalised likelihood simultaneously to find if there is overlap in models. This will determine the models with comparable populations and detected number of events to the LVK data. We check the 30 largest likelihoods out of the 300 new models for both these cases separately, and identify the common models and their hyper-parameters. We again find a clear preference for no CHE for  $f_{WR} \geq 1$  and  $f_\gamma = 0$ , in that these models often return representative numbers of detected events, and populations of masses and redshifts, but



**Figure 12.** Comparison of models with  $\text{SNR} \geq 8$  using the Bayes factor. Each grid is coloured by  $\log_{10}(\mathcal{B})$ . For  $\log_{10}(\mathcal{B}) > 0$ , the model in the row (y-axis) is favoured over than the model in the column (x-axis), while for  $\log_{10}(\mathcal{B}) < 0$  the model in the row is disfavoured over the model in the column. Left plot is the Bayes factor for the marginalised likelihood and right plot is the Bayes factor for the unmarginalised likelihood.



**Figure 13.** Likelihoods for models with no CHE. Each band corresponds to a fixed  $f_T$ . Within each band, excluding the Rauf et al. (2023) models in blue, the mass transfer fraction and CE efficiency are increasing with increasing model number, within the ranges in Table 1, with  $f_{\text{MT}}$  varying more rapidly.

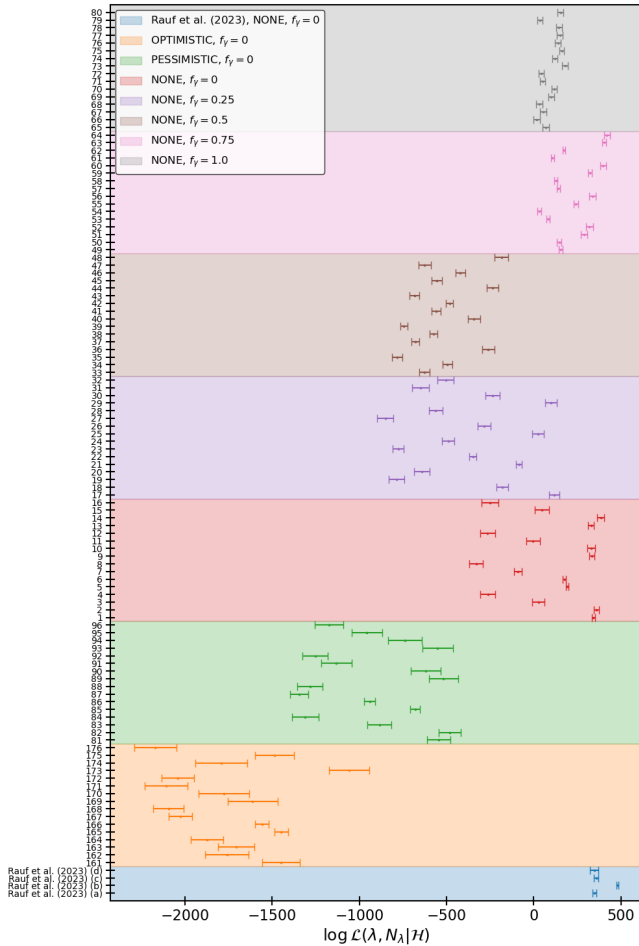


**Figure 14.** Marginalised likelihood for varying CHE and remnant mass prescription. In each band, the Wolf–Rayet multiplier is increased with increasing model number, across the values given in Table 1.

### 5.2.2 Comparisons of the best individual models to data

We now compare the visual population parameter distributions to the GW events with  $\text{SNR} \geq 8$  posteriors for the 300 new models. Model 44 has the best likelihood ( $\mathcal{L}(\lambda|\mathcal{H})$ ) which implies that the model population best matches the observed LVK population. As shown in Fig. 17, we find that the chirp mass and mass ratio distribution falls

the choice of remnant mass prescription varies between Mandel & Müller (2020), Fryer et al. (2012) and Fryer et al. (2022).



**Figure 15.** Model versus unmarginalised likelihood, where each coloured band corresponds to a fixed CHE mode and  $f_\gamma$ . Within each band, excluding the Rauf et al. (2023) models, the mass transfer fraction and CE efficiency are increasing with increasing model number, within the ranges in Table 1, with  $f_{\text{MT}}$  varying more rapidly.

within the  $2\sigma$  contour of the GW events. The LVK chirp mass distribution has three distinct peaks. The best model does not replicate the first peak but it does for the second and third and generally matches the overall chirp mass distribution. That said, the mass ratio distribution is uniform for the LVK data but the model does prefer high mass ratios ( $q > 0.6$ ). Model 79 outputs the worst likelihood. The chirp mass and mass ratio distributions have very narrow peaks corresponding to the second peak in the LVK mass distribution around  $M_c = 20M_\odot$  and at  $q = 0.5$ . The redshift distribution range is also smaller compared to Model 44 and the data. However, the number of detected events,  $\mu_\lambda$  is closer to  $N_{\text{obs}} = 79$ . This demonstrates the intrinsic difficulty in population synthesis of returning the correct event-level properties and overall GW rate.

The best likelihood with a dependence on the number of events ( $\mathcal{L}(\lambda, N_\lambda | \mathcal{H})$ ) is Model 274. This model has a peak in the chirp mass distribution corresponding to the first peak in the LVK mass distribution and recovers the higher mass peaks reasonably well too, albeit with absence in the very high mass tail. The mass ratio distribution is not uniform, there is a preference for mass ratios  $> 0.5$ , but  $< 0.5$  are also possible such that the model and data distributions agree reasonably well. This model outputs  $\mu_\lambda \approx 83$  events. The

worst likelihood is model 199. This model has a rather high peak chirp mass around  $40M_\odot$  and a strong preference for high mass ratios ( $q > 0.8$ ). It is likely that as a result of producing higher mass BHs with smaller separations, this model outputs the largest detected number of events,  $\sim 3805$ , which negatively impacts the likelihood.

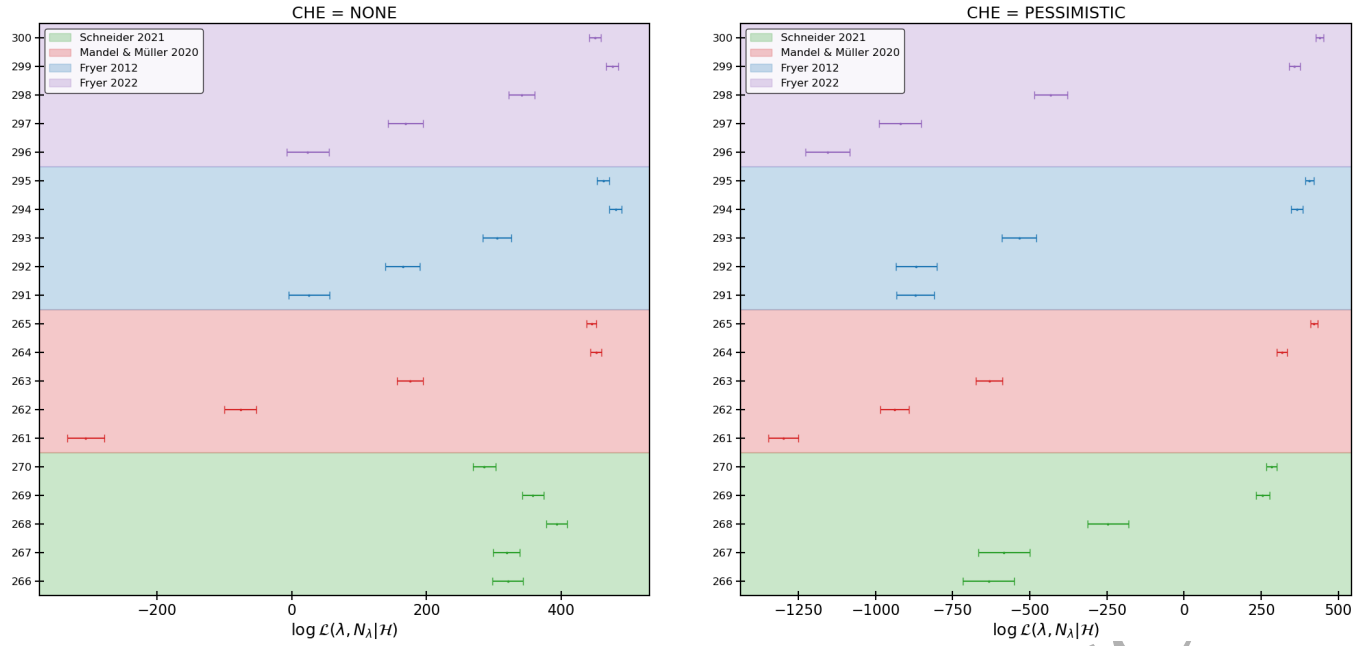
In all cases, our models seem to underpredict the high chirp mass and high mass ratio events with  $z \geq 0.5$  when compared to the LVK distribution. It is possible that these events are not produced frequently by COMPAS. This may be a limitation in the SSE implementation, as discussed in Section 2.4. It is likely that these systems are not produced efficiently through the isolated channel, which could indicate that some of our LVK events progenitors form dynamically. And finally, these events may just be rare or ‘exceptional’ in any formation channel/ environment. In this case, approaches more like Passenger et al. (2024) — which uses the ‘normalised evidence’ within a statistical framework to determine if hierarchical mergers in active galactic nuclei or globular clusters can justify exceptionally large events such as GW190521 — may be more appropriate.

### 5.3 Which data are the most informative?

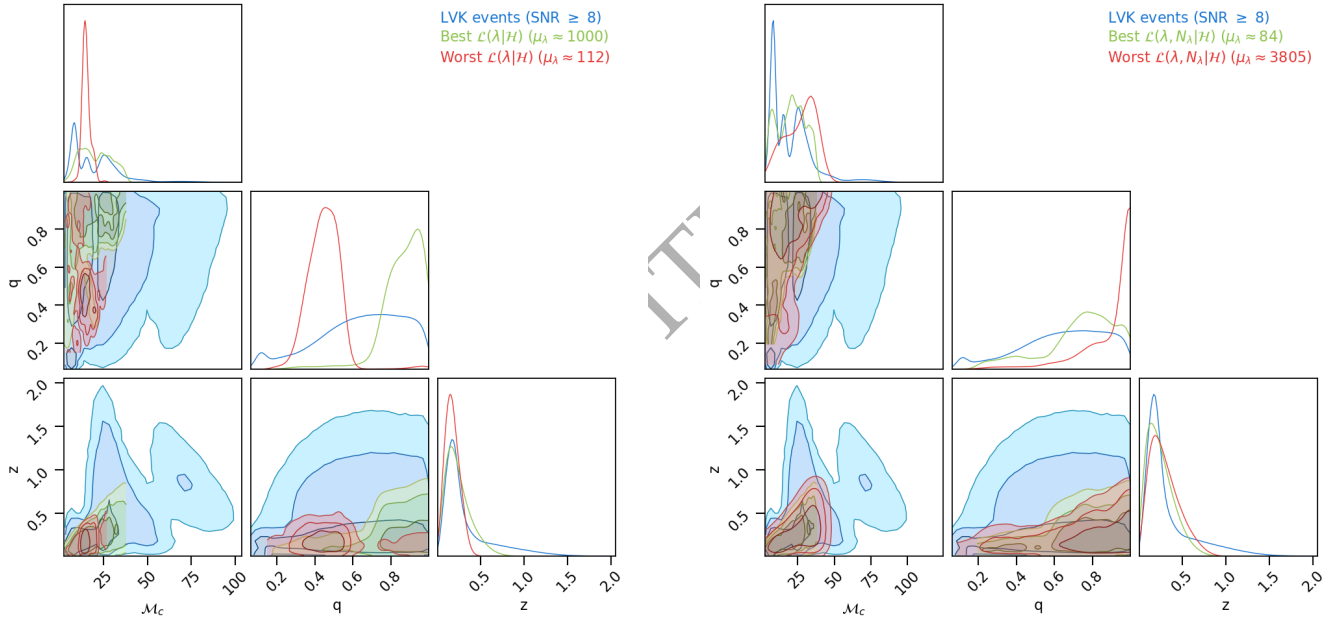
Given our range of models and the form with which the likelihood is evaluated by summing over each GW event, we can also ask the question of which events allow for the greatest distinction between our various COMPAS models. For each GW event, we evaluate the distribution of  $\log(\mathcal{I}^k)$  terms for all 300 models. We use the range and mean of the distributions to determine if the progenitors of the GW events are distinguishable between models. We also look for skewed distributions, which indicate that some models are more or less likely to produce progenitors of particular GW events. We show an example of 20 of these — the top with the least variance between models, and the ten with the most — in Fig. 18.

The top plot of Fig. 18 shows the GW events where their  $\mathcal{I}^k$  terms are tightly constrained to high probabilities. This implies that all the models are very likely to produce the progenitors of these events. The distributions appear Gaussian. However, upon focusing on the  $\log(\mathcal{I}^k) < -5$  region they appear to have a right skewed distribution, implying there are a few models in this tail that are less preferred to form the progenitors of these events.

The bottom plot shows the GW events where the progenitor populations are more clearly distinguishable between models. In particular, GW190814 and GW190521 have wide bimodal distributions, apparently separating our COMPAS models into those that can, and those that cannot, reproduce the characteristics of these events. In Figure 19 we show these events with low probability of their progenitors forming in some models. They appear to be close to, or in, the pair-instability mass gap  $\sim 50 - 130M_\odot$  (Woosley & Heger 2021). GW190426\_190642 has a total mass that supersedes the total mass of GW190521, and it could be argued that they are from the same formation channel. However, GW190426\_190642 is less certain to be a true event, having  $p_{\text{astro}} = 0.75$  and  $\text{FAR} = 4.1\text{yr}^{-1}$  (Banerjee 2022). GW190602\_175927 and GW191109\_010717 have at least one component in the mass gap. The spin-orbit misalignment of GW191109\_010717 provides further evidence it has a dynamical origin (Zhang et al. 2023). However, GW191109\_010717 has been flagged as problematic due to data quality issues, which may lead to unreliable inference of the source parameters such as the precession (Tong et al. 2022). GW190917\_114630 and GW200210\_092254 have a secondary component in the lower mass gap between the most massive neutron star and smallest BH, like GW190814. The low probability of generating these events by some models indicate that the isolated formation channel is not ideal for forming unequal mass



**Figure 16.** Unmarginalised likelihood for No CHE (*Left*) and pessimistic CHE (*Right*) for fixed remnant mass prescriptions. Within each band we vary the Wolf–Rayet multiplier with values given in Table 1, increasing the value with increasing model number.



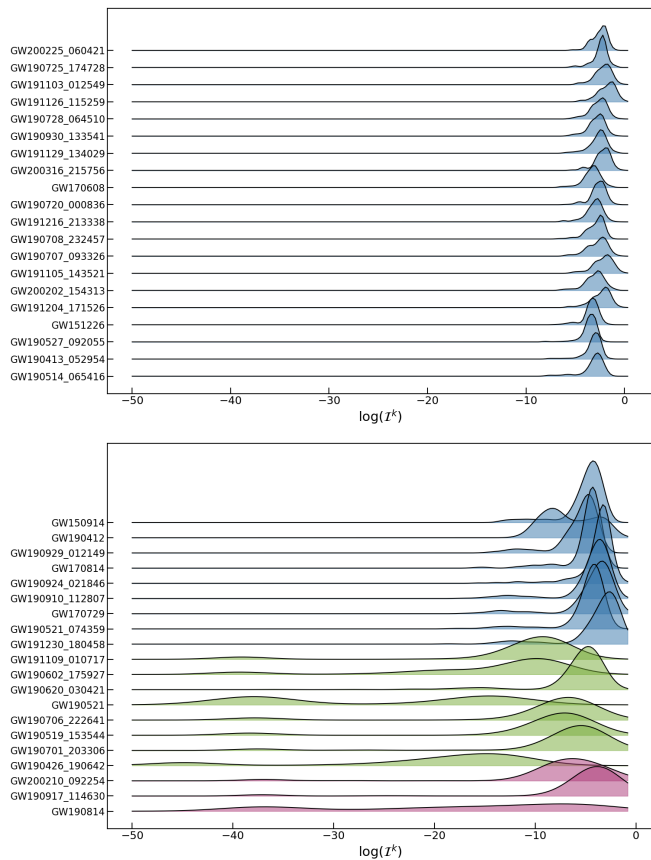
**Figure 17.**  $M_c$ ,  $q$  and  $z$   $3\sigma$  distributions for the models with the best (green contour) and worst (red contour) likelihood. On the left we show the likelihood without the  $N_\lambda$  dependence and the right plot shows the likelihood with the  $N_\lambda$  dependence. The blue contour combines the posteriors for the  $N_{\text{obs}} = 79$  GW events based on the  $\text{SNR} \geq 8$  cut.

binaries. However, the ability of some models to form the progenitors of these events suggests that this mass gap might be narrower or non-existent. More observations of similar events is paramount in confirming this (Zevin et al. 2020).

Overall, looking at the models that perform well for these events, they tend to produce a dearth of higher mass BHs, with chirp mass

distributions that peak around  $40M_\odot$  and mass ratio distributions that peak around  $q \approx 0.9$ . However, this often comes at the expense of producing enough lower mass or more extreme mass ratio binaries, to explain the mass and redshift distribution of the other GW events.

This is particularly emphasised for GW190521 where we find that the posterior samples often did not fall at all within the chirp mass

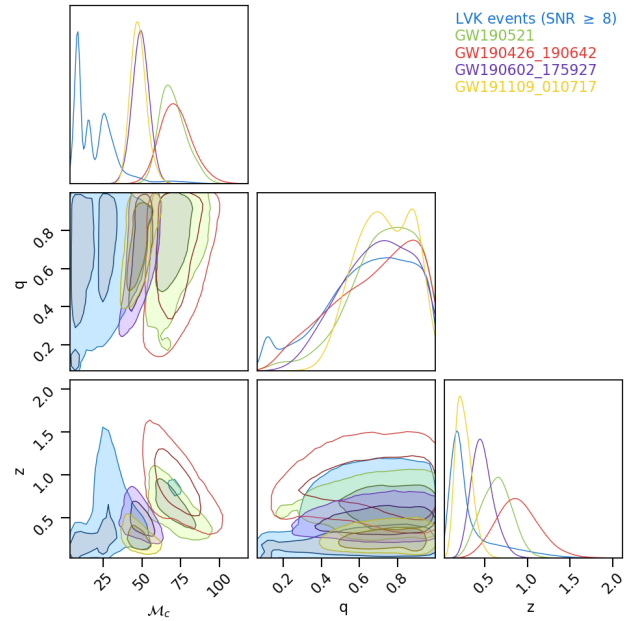


**Figure 18.** Distribution of  $\log(I^k)$  terms for the 300 new COMPAS models. The top plot shows the GW events where the distribution are constrained to high  $\log I^k$  values and the bottom plot shows GW events with broad distributions. Out of the distinguishable events, the pink distributions correspond to events with one component in the lower mass gap and the green distributions correspond to events with at least one component in the upper mass gap.

or mass ratio range output from COMPAS. In these cases we assign a probability of zero to these events such that they make no difference to the sum in Eq. 21 and 26, but are also hence not represented in Fig. 18. Model choices that fail to produce any progenitor close to GW190521 include:

- No CHE (i.e., models 1-80 and 241-250), implying that formation channels with mass transfer and CE evolution produce smaller BH masses.
- The remnant mass prescriptions Mandel & Müller (2020) and Schneider et al. (2021), which tend to produce smaller BH masses. The exception being model 271-280, which use Fryer et al. (2012) and Fryer et al. (2021).
- Models 221 and 270, where CHE is set to Optimistic, but  $f_\gamma = 0.75$  (Model 221) and/or  $f_{WR} = 10$  (Model 270), which may be contributing to extreme mass loss from the progenitors again leading to a population of particularly low-mass BHs.

Overall, the difficulty in simultaneously producing both these larger events and the bulk of LVK data even across the large range of models we explore in this work lends credence to the idea of multiple formation pathways for these events.



**Figure 19.** Posterior distributions for ‘outlier’ events, (identified by the green distributions in Fig. 18) that are not well reproduced by the majority of our COMPAS models. The blue contour is the full posterior distribution for LVK data with  $\text{SNR} \geq 8$ , which clearly indicates the majority of the data are at both lower mass and lower redshift. We show GW190426\_190642 as an unfilled contour, as it does not fit the  $p_{\text{astro}} \geq 0.9$  and  $\text{FAR} \leq 0.25\text{yr}^{-1}$  criteria.

## 6 CONCLUSION

In this work we present a thorough analysis of the ability of over 300 binary population synthesis models to reproduce the GW events observed so far within Gravitational Wave Transients Catalogues 1-3. To enable this, we developed an efficient method for predicting the BBH merger rate in the isolated channel for various COMPAS models (Section 3), validated based on previous work from Rauf et al. (2023). We then used this within a Bayesian inference method to calculate the likelihood of these models with and without the dependence on the number of GW events predicted by each model,  $N_\lambda$  (Section 4).

The COMPAS population synthesis code contains a number of free parameters — we introduce some of the most salient of these in Section 2. These are the mass transfer fraction, CE efficiency, specific AM loss fraction, CHE, remnant mass prescription and Wolf-Rayet winds. We find that various combinations of these hyper-parameters impact the chirp mass, mass ratio and redshift distribution, as well as changing the orbital separations of binaries, which affects the number of BBH mergers within an observed time. Our results fitting these models to data in Section 5 demonstrate there is a preference for subsets of models dictated by the type of isolated formation channel and specific AM loss. However, we frequently find opposing trends in the likelihoods with and without  $N_\lambda$  dependence — models that produce realistic event-level masses and redshifts frequently over-predict the number of events we should have observed to date. Looking at models which perform well with both likelihoods, we find that models with no CHE, low to moderate values of  $f_\gamma$  (such that AM loss occurs closer to the donor) and higher rates of mass-loss from Wolf-Rayet winds ( $f_{WR} > 1$ ) are generally preferred by current data. We provide a quantitative summary of the hyper-parameters for the models that best reproduce the GW population excluding GW190521 at the end of Section 5.3.

We find that the selection effects impact our likelihoods. Although the selection allows for potentially non-astrophysical sources in the sample, the  $\text{SNR} \geq 8$  cut on the GW events increases the likelihood for *all* models compared to  $\text{FAR} \leq 0.25\text{yr}^{-1}$  and  $p_{\text{astro}} \geq 0.9$ . Preference for models do not change with either selection criteria. However, these conclusions may vary with future GW detections, particularly if the BH masses are in the mass gaps. We also use the individual event likelihoods  $\mathcal{I}^k$  across all our models term to distinguish individual events that are easy or more difficult to reproduce by our models. We identify the usual suspects GW190521 and GW190814 as typically having low probability of existing in our new models, as well as some outlier events in the the lower and upper mass gap. This motivates future studies to consider joint population inference and model selection studies across multiple formation channels (Gompertz et al. 2022; Gayathri et al. 2023).

Our method for efficiently calculating the merger rate can also be applied to other population synthesis codes. However, we note that it is only applicable in the case where the metallicity dependent SFR is the same for the fiducial and new model. We argue that using a different SFR may not impact the PDF significantly. van Son et al. (2023) finds that variation in the SFR do not impact the features in the primary BH mass distribution and have minimal impact on the low mass end of the distribution. For the redshift range studied in this paper, SHARK predicts a metallicity dependent SFR density compared to the fiducial model explored by Neijssel et al. (2019), which produces a BBH merger rate comparable to GWTC-1. This may be contributing to our models overestimating the number of events. Variations in stellar evolution in population synthesis (Broekgaarden et al. 2022a; Chruślińska 2024) and modelling of galaxy distributions in galaxy simulations<sup>11</sup> (Furlong et al. 2015; Lagos et al. 2024) can translate to large uncertainties in the SFR, that will impact the BBH merger rate. These uncertainties are non-negligible and should be accounted for in future studies of BBH model selection.

We acknowledge this work relies on various assumptions and simplifications. For future work, we consider including the effective spin as a population parameter in the analysis. The covariance between the mass ratio and effective spin can arise due to CE physics, mass transfer efficiency and Eddington limited accretion (Mandel & Fragos 2020; Bavera et al. 2021; Zevin & Bavera 2022; Broekgaarden et al. 2022b; Adamcewicz et al. 2023; Olejak et al. 2024). This could be used to further distinguish between COMPAS models. Alternatively, we can build the PDF only with  $M_c$ , mass ratio or redshift to test the validity of the model outputs. We could also consider including natal kick velocity as one of the hyper-parameters to vary in COMPAS in the future, as this is not well constrained for BHs (Callister et al. 2021; Cui & Li 2023; Vigna-Gómez et al. 2024), or more fundamental parameters such as the binary fraction. Varying the selection function cuts applied to current data to restrict the analysis to only the strongest detections may further inform us on trends but given we find consistent results with our two different selection functions this may not necessarily provide improved constraints on the hyper-parameters.

Future GW detectors will observe GW events at a wider range of frequencies and redshifts, and will provide data with increasing precision (Broekgaarden et al. 2023). This will answer a variety of astrophysical questions and provide a clearer picture of the BBH population distribution. In the absence of electromagnetic counterparts, BBH merger populations with the *correct* modelling are crucial for an unbiased inference of cosmological parameters (Mastrogiovanni

et al. 2021; Gray et al. 2023). With deeper catalogues and higher SNR detections, our selection criteria will be more refined, which will hopefully allow us to further constrain the uncertainties in massive binary evolution models. However, as always, with more GW events comes greater change of observing rare exceptional events to challenge these models. Until then, in the case of limited number of GW events we must rely on observations of massive binary stars to constrain the different assumptions for massive binary evolution. Observations of X-ray binaries, red supergiants, luminous blue variables and Wolf–Rayet stars can help constrain binaries at different evolutionary stages (Eldridge & Stanway 2022; Marchant & Bodensteiner 2023; Li et al. 2024).

## ACKNOWLEDGEMENTS

We thank Carole Perigois and Michella Mapelli for useful discussion on their paper that motivated this work. We thank Ilya Mandel, Eric Thrane and Paul Lasky for useful discussions on the methodology and results. We also thank Floor Broekgaarden for useful discussions on the COMPAS hyper-parameters and relevant literature. We also thank the anonymous reviewer for their helpful comments, that improved the quality of the manuscript.

The authors acknowledge support from the Australian Research Council (ARC) Centre of Excellence for Gravitational Wave Discovery (OzGrav), through project numbers CE170100004 and CE230100016. LR and CH acknowledge support through ARC Discovery Project DP20220101395. SS is a recipient of an ARC Discovery Early Career Research Award (DE220100241).

## DATA AVAILABILITY

Prior and posterior samples of GW events were provided by Carole Perigois. Simulations in this paper made use of the COMPAS rapid binary population synthesis code (version v02.38.08), which is freely available at [this repository](#).

## REFERENCES

- Abbott B. P., et al., 2017, *Phys. Rev. Lett.*, **119**, 161101  
 Abbott B. P., et al., 2019a, *Physical Review X*, **9**, 031040  
 Abbott B. P., et al., 2019b, *ApJ*, **882**, L24  
 Abbott R., et al., 2020a, *Phys. Rev. Lett.*, **125**, 101102  
 Abbott B. P., et al., 2020b, *ApJ*, **892**, L3  
 Abbott R., et al., 2020c, *ApJ*, **896**, L44  
 Abbott R., et al., 2021a, *Physical Review X*, **11**, 021053  
 Abbott R., et al., 2021b, *ApJ*, **915**, L5  
 Abbott R., et al., 2023a, *Physical Review X*, **13**, 011048  
 Abbott R., et al., 2023b, *Physical Review X*, **13**, 041039  
 Adamcewicz C., Lasky P. D., Thrane E., 2023, *ApJ*, **958**, 13  
 Agrawal P., Hurley J., Stevenson S., Szécsi D., Flynn C., 2020, *MNRAS*, **497**, 4549  
 Agrawal P., Szécsi D., Stevenson S., Eldridge J. J., Hurley J., 2022a, *MNRAS*, **512**, 5717  
 Agrawal P., Stevenson S., Szécsi D., Hurley J., 2022b, *A&A*, **668**, A90  
 Andres N., et al., 2022, *Classical and Quantum Gravity*, **39**, 055002  
 Arca Sedda M., 2021, *ApJ*, **908**, L38  
 Arca Sedda M., Mapelli M., Spera M., Benacquista M., Giacobbo N., 2020, *ApJ*, **894**, 133  
 Arca Sedda M., Mapelli M., Benacquista M., Spera M., 2023, *MNRAS*, **520**, 5259  
 Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, **47**, 481

<sup>11</sup> We refer to Fig. 2 and 3 of Rauf et al. (2023) for comparisons of SHARK to observations.

- Banagiri S., Berry C. P. L., Davies G. S. C., Tsukada L., Doctor Z., 2023, *Phys. Rev. D*, **108**, 083043
- Banerjee S., 2022, *A&A*, **665**, A20
- Barkat Z., Rakavy G., Sack N., 1967, *Phys. Rev. Lett.*, **18**, 379
- Barrett J. W., Gaebel S. M., Neijssel C. J., Vigna-Gómez A., Stevenson S., Berry C. P. L., Farr W. M., Mandel I., 2018, *MNRAS*, **477**, 4685
- Bavera S. S., et al., 2021, *A&A*, **647**, A153
- Belczynski K., Bulik T., Fryer C. L., Ruiter A., Valsecchi F., Vink J. S., Hurley J. R., 2010, *ApJ*, **714**, 1217
- Björklund R., Sundqvist J. O., Singh S. M., Puls J., Najjaro F., 2023, *A&A*, **676**, A109
- Bogomazov A. I., Cherepashchuk A. M., Lipunov V. M., Tutukov A. V., 2018, *New Astron.*, **58**, 33
- Bouffanais Y., Mapelli M., Gerosa D., Di Carlo U. N., Giacobbo N., Berti E., Baibhav V., 2019, *ApJ*, **886**, 25
- Bouffanais Y., Mapelli M., Santoliquido F., Giacobbo N., Iorio G., Costa G., 2021, *MNRAS*, **505**, 3873
- Breivik K., et al., 2020a, *ApJ*, **898**, 71
- Breivik K., et al., 2020b, *ApJ*, **898**, 71
- Broekgaarden F. S., Berger E., 2021, *ApJ*, **920**, L13
- Broekgaarden F. S., et al., 2021, *MNRAS*, **508**, 5028
- Broekgaarden F. S., et al., 2022a, *MNRAS*, **516**, 5737
- Broekgaarden F. S., Stevenson S., Thrane E., 2022b, *ApJ*, **938**, 45
- Broekgaarden F. S., Banagiri S., Payne E., 2023, *arXiv e-prints*, p. [arXiv:2303.17628](https://arxiv.org/abs/2303.17628)
- Bulik T., Belczynski K., 2010, *Mem. Soc. Astron. Italiana*, **81**, 302
- Cabourn Davies G. S., Harry I. W., 2022, *Classical and Quantum Gravity*, **39**, 215012
- Callister T. A., Kanner J. B., Massinger T. J., Dhurandhar S., Weinstein A. J., 2017, *Classical and Quantum Gravity*, **34**, 155007
- Callister T. A., Farr W. M., Renzo M., 2021, *ApJ*, **920**, 157
- Cheng A. Q., Zevin M., Vitale S., 2023a, *ApJ*, **955**, 127
- Cheng A. Q., Zevin M., Vitale S., 2023b, *Astrophys. J.*, **955**, 127
- Chruślińska M., 2024, *Annalen der Physik*, **536**, 2200170
- Chu Q., Yu S., Lu Y., 2022, *MNRAS*, **509**, 1557
- Claret A., Torres G., 2018, *ApJ*, **859**, 100
- Costa G., Girardi L., Bressan A., Marigo P., Rodrigues T. S., Chen Y., Lanza A., Goudfrooij P., 2019, *MNRAS*, **485**, 4641
- Costa G., Chruślińska M., Klencki J., Broekgaarden F. S., Rodriguez C. L., Joseph T. D., Saracino S., 2023, *arXiv e-prints*, p. [arXiv:2311.15778](https://arxiv.org/abs/2311.15778)
- Cui Z., Li X.-D., 2023, *MNRAS*, **523**, 5565
- Delfavero V., O’Shaughnessy R., Belczynski K., Drozda P., Wysocki D., 2023, *Phys. Rev. D*, **108**, 043023
- Deng Z.-L., Li X.-D., Shao Y., Xu K., 2024, *ApJ*, **963**, 80
- Detmers R. G., Langer N., Podsiadlowski P., Izzard R. G., 2008, *A&A*, **484**, 831
- Dorozsmai A., Toonen S., 2022, *arXiv e-prints*, p. [arXiv:2207.08837](https://arxiv.org/abs/2207.08837)
- Eldridge J. J., Stanway E. R., 2022, *ARA&A*, **60**, 455
- Ertl T., Janka H. T., Woosley S. E., Sukhbold T., Ugliano M., 2016, *ApJ*, **818**, 124
- Farr W. M., Gair J. R., Mandel I., Cutler C., 2015, *Phys. Rev. D*, **91**, 023005
- Fishbach M., Kalogera V., 2022, *ApJ*, **929**, L26
- Fishbach M., Holz D. E., Farr W. M., 2018, *ApJ*, **863**, L41
- Fowler W. A., Hoyle F., 1964, *ApJS*, **9**, 201
- Fragos T., Andrews J. J., Ramirez-Ruiz E., Meynet G., Kalogera V., Taam R. E., Zezas A., 2019, *ApJ*, **883**, L45
- Fryer C., 2015, *PoS, FRAPWS2014*, 004
- Fryer C. L., Belczynski K., Wiktorowicz G., Dominik M., Kalogera V., Holz D. E., 2012, *The Astrophysical Journal*, **749**, 91
- Fryer C. L., Karpov P., Livescu D., 2021, *Astronomy Reports*, **65**, 937
- Fryer C. L., Olejak A., Belczynski K., 2022, *ApJ*, **931**, 94
- Furlong M., et al., 2015, *MNRAS*, **450**, 4486
- Galatgade S., Talbot C., Thrane E., 2020, *Phys. Rev. D*, **102**, 083026
- Gallegos-García M., Berry C. P. L., Marchant P., Kalogera V., 2021, *ApJ*, **922**, 110
- Gallegos-García M., Jacquemin-Ide J., Kalogera V., 2023, *arXiv e-prints*, p. [arXiv:2308.13146](https://arxiv.org/abs/2308.13146)
- García F., Simaz Bunzel A., Chaty S., Porter E., Chassande-Mottin E., 2021, *A&A*, **649**, A114
- Gayathri V., Bartos I., Rosswog S., Miller M. C., Veske D., Lu W., Marka S., 2023, *arXiv e-prints*, p. [arXiv:2307.09097](https://arxiv.org/abs/2307.09097)
- Ge H., Hjellming M. S., Webbink R. F., Chen X., Han Z., 2010, *ApJ*, **717**, 724
- Ge H., Webbink R. F., Chen X., Han Z., 2015, *ApJ*, **812**, 40
- Ge H., Webbink R. F., Chen X., Han Z., 2020, *ApJ*, **899**, 132
- Gompertz B. P., Nicholl M., Schmidt P., Pratten G., Vecchio A., 2022, *MNRAS*, **511**, 1454
- Gray R., et al., 2023, *J. Cosmology Astropart. Phys.*, **2023**, 023
- Grichener A., 2023, *MNRAS*, **523**, 221
- Han Z.-W., Ge H.-W., Chen X.-F., Chen H.-L., 2020, *Research in Astronomy and Astrophysics*, **20**, 161
- Hirai R., Mandel I., 2022, *ApJ*, **937**, L42
- Hurley J. R., Pols O. R., Tout C. A., 2000, *MNRAS*, **315**, 543
- Hurley J. R., Tout C. A., Pols O. R., 2002, *MNRAS*, **329**, 897
- Ivanova N., et al., 2013, *A&ARv*, **21**, 59
- Ivanova N., Justham S., Podsiadlowski P., 2015, *MNRAS*, **447**, 2181
- Janka H.-T., 2012, *Annual Review of Nuclear and Particle Science*, **62**, 407
- Janka H.-T., 2017, in Alsabti A. W., Murdin P., eds., *Handbook of Supernovae*. p. 1095, doi:[10.1007/978-3-319-21846-5\\_109](https://doi.org/10.1007/978-3-319-21846-5_109)
- Janka H. T., Langanke K., Marek A., Martínez-Pinedo G., Müller B., 2007, *Phys. Rep.*, **442**, 38
- Kaaret P., Alonso-Herrero A., Gallagher J. S., Fabbiano G., Zezas A., Rieke M. J., 2004, *MNRAS*, **348**, L28
- Klencki J., Nelemans G., Istrate A. G., Chruslińska M., 2021, *A&A*, **645**, A54
- Kowalska-Leszczynska I., Regimbau T., Bulik T., Dominik M., Belczynski K., 2015, *A&A*, **574**, A58
- Kroupa P., 2002, *Science*, **295**, 82
- LIGO Scientific Collaboration Virgo Collaboration KAGRA Collaboration 2018, LVK Algorithm Library - LALSuite, Free software (GPL), doi:[10.7935/GT1W-FZ16](https://doi.org/10.7935/GT1W-FZ16)
- Lagos C. d. P., Tobar R. J., Robotham A. S. G., Obreschkow D., Mitchell P. D., Power C., Elahi P. J., 2018, *MNRAS*, **481**, 3573
- Lagos C. d. P., et al., 2019, *MNRAS*, **489**, 4196
- Lagos C. d. P., et al., 2024, *MNRAS*, **531**, 3551
- Laplace E., Göteborg Y., de Mink S. E., Justham S., Farmer R., 2020, *A&A*, **637**, A6
- Lau M. Y. M., Hirai R., Price D. J., Mandel I., 2022, *MNRAS*, **516**, 4669
- Li Z., Zhu C., Lü G., Li L., Liu H., Guo S., Yu J., Lu X., 2024, *arXiv e-prints*, p. [arXiv:2405.11571](https://arxiv.org/abs/2405.11571)
- Liotine C., Zevin M., Berry C. P. L., Doctor Z., Kalogera V., 2023, *ApJ*, **946**, 4
- Lu W., Fuller J., Quataert E., Bonnerot C., 2023, *MNRAS*, **519**, 1409
- Lynch R., Coughlin M., Vitale S., Stubbs C. W., Katsavounidis E., 2018, *ApJ*, **861**, L24
- Lynden-Bell D., Wood R., 1968, *MNRAS*, **138**, 495
- MacLeod M., Loeb A., 2020, *ApJ*, **895**, 29
- MacLeod M., Ostriker E. C., Stone J. M., 2018a, *ApJ*, **863**, 5
- MacLeod M., Ostriker E. C., Stone J. M., 2018b, *ApJ*, **868**, 136
- Mandel I., Broekgaarden F. S., 2021, *arXiv preprint arXiv:2107.14239*
- Mandel I., Broekgaarden F. S., 2022, *Living Reviews in Relativity*, **25**, 1
- Mandel I., Farmer A., 2022, *Phys. Rep.*, **955**, 1
- Mandel I., Fragos T., 2020, *ApJ*, **895**, L28
- Mandel I., Müller B., 2020, *Monthly Notices of the Royal Astronomical Society*, **499**, 3214
- Mandel I., de Mink S. E., 2016, *MNRAS*, **458**, 2634
- Mandel I., Farr W. M., Gair J. R., 2019, *MNRAS*, **486**, 1086
- Mapelli M., 2019, in Oskinova L. M., Bozzo E., Bulik T., Gies D. R., eds., *Vol. 346, High-mass X-ray Binaries: Illuminating the Passage from Massive Binaries to Merging Compact Objects*. pp 397–416, doi:[10.1017/S1743921319001868](https://doi.org/10.1017/S1743921319001868)
- Mapelli M., 2020, *Frontiers in Astronomy and Space Sciences*, **7**, 38
- Mapelli M., 2021, in *Handbook of Gravitational Wave Astronomy*. p. 16, doi:[10.1007/978-981-15-4702-7\\_16-1](https://doi.org/10.1007/978-981-15-4702-7_16-1)
- Mapelli M., Giacobbo N., 2018, *MNRAS*, **479**, 4391

- Marchant P., Bodensteiner J., 2023, [arXiv e-prints](#), p. [arXiv:2311.01865](#)
- Marchant P., Langer N., Podsiadlowski P., Tauris T. M., Moriya T. J., 2016, *A&A*, **588**, A50
- Marchant P., Pappas K. M. W., Gallegos-Garcia M., Berry C. P. L., Taam R. E., Kalogera V., Podsiadlowski P., 2021, *A&A*, **650**, A107
- Marchant P., Podsiadlowski P., Mandel I., 2023, [arXiv e-prints](#), p. [arXiv:2311.14041](#)
- Marek A., Janka H. T., 2009, *ApJ*, **694**, 664
- Mastrogiovanni S., et al., 2021, *Phys. Rev. D*, **104**, 062009
- Mennekens N., Vanbeveren D., 2014, *A&A*, **564**, A134
- Miller-Jones J. C. A., et al., 2021, *Science*, **371**, 1046
- Moreno Méndez E., López-Cámara D., De Colle F., 2017, *MNRAS*, **470**, 2929
- Müller B., Heger A., Liptai D., Cameron J. B., 2016, *MNRAS*, **460**, 742
- Nathaniel K., 2017, Modeling Chemically Homogeneous Evolution in COMPAS, [https://www.phys.ufl.edu/ireu/IREU2017/pdf\\_reports/Kaila\\_Nathaniel\\_Final\\_Report\\_BHam.pdf](https://www.phys.ufl.edu/ireu/IREU2017/pdf_reports/Kaila_Nathaniel_Final_Report_BHam.pdf)
- Neijssel C. J., et al., 2019, *MNRAS*, **490**, 3740
- Neijssel C. J., Vinciguerra S., Vigna-Gómez A., Hirai R., Miller-Jones J. C. A., Bahramian A., Maccarone T. J., Mandel I., 2021, *ApJ*, **908**, 118
- Neo S., Miyaji S., Nomoto K., Sugimoto D., 1977, *PASJ*, **29**, 249
- O'Connor E., Ott C. D., 2011, *ApJ*, **730**, 70
- Ohlmann S. T., Röpke F. K., Pakmor R., Springel V., 2016, *ApJ*, **816**, L9
- Olejak A., Fryer C. L., Belczynski K., Baibhav V., 2022, *MNRAS*, **516**, 2252
- Olejak A., Klencik J., Xu X.-T., Wang C., Belczynski K., Lasota J.-P., 2024, [arXiv e-prints](#), p. [arXiv:2404.12426](#)
- Orosz J. A., et al., 2007, *Nature*, **449**, 872
- Paczynski B., Sienkiewicz R., 1972, *Acta Astron.*, **22**, 73
- Passenger L., Thrane E., Lasky P. D., Payne E., Stevenson S., Farr B., 2024, [arXiv e-prints](#), p. [arXiv:2405.09739](#)
- Paxton B., Bildsten L., Dotter A., Herwig F., Lesaffre P., Timmes F., 2011a, *ApJS*, **192**, 3
- Paxton B., Bildsten L., Dotter A., Herwig F., Lesaffre P., Timmes F., 2011b, *ApJS*, **192**, 3
- Paxton B., et al., 2013, *ApJS*, **208**, 4
- Paxton B., et al., 2015, *ApJS*, **220**, 15
- Paxton B., et al., 2018, *ApJS*, **234**, 34
- Paxton B., et al., 2019, *ApJS*, **243**, 10
- Pérido C., Mapelli M., Santoliquido F., Bouffanais Y., Rufolo R., 2023, [arXiv e-prints](#), p. [arXiv:2301.01312](#)
- Picco A., Marchant P., Sana H., Nelemans G., 2023, [arXiv e-prints](#), p. [arXiv:2309.05736](#)
- Picker L., Hirai R., Mandel I., 2024, [arXiv e-prints](#), p. [arXiv:2402.13180](#)
- Pols O. R., Schröder K.-P., Hurley J. R., Tout C. A., Eggleton P. P., 1998, *MNRAS*, **298**, 525
- Powell J., Gossan S. E., Logue J., Heng I. S., 2016, *Phys. Rev. D*, **94**, 123012
- Powell J., Iess A., Llorens-Monteagudo M., Obergaullinger M., Müller B., Torres-Forné A., Cuoco E., Font J. A., 2023, [arXiv e-prints](#), p. [arXiv:2311.18221](#)
- Rauf L., Howlett C., Davis T. M., Lagos C. D. P., 2023, *Monthly Notices of the Royal Astronomical Society*, **523**, 5719
- Riley J., Mandel I., 2023, *ApJ*, **950**, 80
- Riley J., Mandel I., Marchant P., Butler E., Nathaniel K., Neijssel C., Shortt S., Vigna-Gómez A., 2021, *MNRAS*, **505**, 663
- Rodriguez C. L., Zevin M., Amaro-Seoane P., Chatterjee S., Kremer K., Rasio F. A., Ye C. S., 2019, *Phys. Rev. D*, **100**, 043027
- Romagnolo A., Belczynski K., Klencik J., Agrawal P., Shenar T., Szécsi D., 2023, *MNRAS*, **525**, 706
- Romero-Shaw I., Lasky P. D., Thrane E., 2022, *ApJ*, **940**, 171
- Romero-Shaw I., Hirai R., Bahramian A., Willcox R., Mandel I., 2023, *MNRAS*, **524**, 245
- Röpke F. K., De Marco O., 2023, *Living Reviews in Computational Astrophysics*, **9**, 2
- Sana H., et al., 2012, *Science*, **337**, 444
- Santoliquido F., Mapelli M., Giacobbo N., Bouffanais Y., Artale M. C., 2021, *MNRAS*, **502**, 4877
- Santoliquido F., Mapelli M., Iorio G., Costa G., Glover S. C. O., Hartwig T., Klessen R. S., Merli L., 2024, *MNRAS*, **528**, 954
- Schneider F. R. N., Izzard R. G., Langer N., de Mink S. E., 2015, *ApJ*, **805**, 20
- Schneider F., Podsiadlowski P., Müller B., 2021, *Astronomy & Astrophysics*, **645**, A5
- Schneider F. R. N., Podsiadlowski P., Laplace E., 2023, *ApJ*, **950**, L9
- Schneider F. R. N., Podsiadlowski P., Laplace E., 2024, [arXiv e-prints](#), p. [arXiv:2403.03984](#)
- Schröder S. L., MacLeod M., Loeb A., Vigna-Gómez A., Mandel I., 2020, *ApJ*, **892**, 13
- Schröder S. L., MacLeod M., Ramirez-Ruiz E., Mandel I., Fragos T., Loeb A., Everson R. W., 2021, [arXiv e-prints](#), p. [arXiv:2107.09675](#)
- Shenar T., Gilkis A., Vink J. S., Sana H., Sander A. A. C., 2020, *A&A*, **634**, A79
- Silverman B. W., 1986, Density estimation for statistics and data analysis
- Spera M., Mapelli M., Bressan A., 2015, *MNRAS*, **451**, 4086
- Stevenson S., Clarke T. A., 2022, *MNRAS*, **517**, 4034
- Stevenson S., Ohme F., Fairhurst S., 2015, *ApJ*, **810**, 58
- Stevenson S., Vigna-Gómez A., Mandel I., Barrett J. W., Neijssel C. J., Perkins D., de Mink S. E., 2017, *Nature Communications*, **8**, 14906
- Tanikawa A., 2024, [arXiv e-prints](#), p. [arXiv:2403.04389](#)
- Team COMPAS: Riley, J. et al., 2022, *ApJS*, **258**, 34
- Temaj D., Schneider F. R. N., Laplace E., Wei D., Podsiadlowski P., 2024, *A&A*, **682**, A123
- Temminck K. D., Pols O. R., Justham S., Istrate A. G., Toonen S., 2023, *A&A*, **669**, A45
- Tong H., Galadage S., Thrane E., 2022, *Phys. Rev. D*, **106**, 103019
- Tutukov A., Yungelson L., 1979, in Conti P. S., De Loore C. W. H., eds., Vol. 83, Mass Loss and Evolution of O-Type Stars. pp 401–406
- Vigna-Gómez A., et al., 2018, *MNRAS*, **481**, 4009
- Vigna-Gómez A., et al., 2024, *Phys. Rev. Lett.*, **132**, 191403
- Vink J. S., Harries T. J., 2017, *A&A*, **603**, A120
- Vink J. S., de Koter A., 2005, *A&A*, **442**, 587
- Vitale S., Gerosa D., Farr W. M., Taylor S. R., 2022, in , Handbook of Gravitational Wave Astronomy. p. 45, doi:10.1007/978-981-15-4702-7\_45-1
- Webbink R. F., 1984, *ApJ*, **277**, 355
- Willcox R., MacLeod M., Mandel I., Hirai R., 2023, *ApJ*, **958**, 138
- Winch E. R. J., Vink J. S., Higgins E. R., Sabhahitf G. N., 2024, *MNRAS*, **529**, 2980
- Wong K. W. K., Breivik K., Farr W. M., Luger R., 2023, *ApJ*, **950**, 181
- Woosley S. E., Heger A., 2021, *ApJ*, **912**, L31
- Xu X.-J., Li X.-D., 2010, *ApJ*, **716**, 114
- Yang Y., Bartos I., Haiman Z., Kocsis B., Márka Z., Stone N. C., Márka S., 2019, *ApJ*, **876**, 122
- Zevin M., Bavera S. S., 2022, *ApJ*, **933**, 86
- Zevin M., Spera M., Berry C. P. L., Kalogera V., 2020, *ApJ*, **899**, L1
- Zevin M., et al., 2021a, *ApJ*, **910**, 152
- Zevin M., Romero-Shaw I. M., Kremer K., Thrane E., Lasky P. D., 2021b, *ApJ*, **921**, L43
- Zhang R. C., Fragione G., Kimball C., Kalogera V., 2023, *ApJ*, **954**, 23
- Zheng Y., Cavaglia M., Quitzow-James R., Mogushi K., 2021, *Significance*, **18**
- Zuo Z.-Y., Li X.-D., 2014, *MNRAS*, **442**, 1980
- de Kool M., 1990, *ApJ*, **358**, 189
- de Mink S. E., Pols O. R., Hilditch R. W., 2007, *A&A*, **467**, 1181
- de Mink S. E., Cantiello M., Langer N., Pols O. R., Brott I., Yoon S. C., 2009, *A&A*, **497**, 243
- de Mink S. E., Cantiello M., Langer N., Pols O. R., Yoon S. C., 2010, in Prša A., Zejda M., eds, Astronomical Society of the Pacific Conference Series Vol. 435, Binaries - Key to Comprehension of the Universe. p. 179 ([arXiv:0910.3694](#)), doi:10.48550/arXiv.0910.3694
- du Buisson L., et al., 2020, *MNRAS*, **499**, 5941
- van Son L. A. C., et al., 2020, *ApJ*, **897**, 100
- van Son L. A. C., et al., 2022a, *ApJ*, **931**, 17
- van Son L. A. C., et al., 2022b, *ApJ*, **940**, 184
- van Son L. A. C., de Mink S. E., Chruślińska M., Conroy C., Pakmor R., Hernquist L., 2023, *ApJ*, **948**, 105
- van den Heuvel E. P. J., 1976, in Eggleton P., Mitton S., Whelan J., eds, IAU

Symposium Vol. 73, Structure and Evolution of Close Binary Systems.  
p. 35

van den Heuvel E. P. J., Portegies Zwart S. F., de Mink S. E., 2017, *MNRAS*,  
471, 4256

#### **APPENDIX A: MODEL INITIAL CONDITIONS AND FIXED PARAMETERS**

In Table A1, we list the distributions and ranges for the input parameters used to define the initial conditions of our COMPAS binaries. These are fixed for all models explored in this paper and Rauf et al. (2023).

#### **APPENDIX B: GW DATA**

In Table B1, we list the GW events used as data for each selection criteria, as described in Section 4.1.

#### **APPENDIX C: MODEL HYPER-PARAMETERS AND BAYES FACTOR DATA**

Extending on Table 1, for each model in Tables C1, C2 and C3 we detail the hyper-parameters and Bayes factor for each selection criteria. Rauf et al. (2023) (a)-(d) refer to the models explored in Rauf et al. (2023), where we vary the remnant mass prescription (with the exception of Rauf et al. (2023) (c) where we also vary  $f_{\text{WR}}$ ).

ORIGINAL UNEDITED MANUSCRIPT

Distribution	Model	Range
IMF	<a href="#">Kroupa (2002)</a>	$[5, 150]M_{\odot}$ $[5, 200]M_{\odot}$ ( <a href="#">Rauf et al. 2023</a> )
Semi-major axis	Flat-in-log	$[0.01, 1000]$ AU
Orbital period	Flat-in-log	$[0.1, 1000]$ days
Mass ratio	<a href="#">Sana et al. (2012)</a>	$[0.01, 1]$
Eccentricity	ZERO (circular) <a href="#">Sana et al. (2012)</a> ( <a href="#">Rauf et al. 2023</a> )	0 $[0, 1]$
Log Metallicity	SHARK ( <a href="#">Lagos et al. 2018</a> )	$[-7.0, -0.3]$

**Table A1.** Summary of COMPAS default initial conditions and range.

ORIGINAL UNEDITED MANUSCRIPT

GW events	$FAR \leq 0.25\text{yr}^{-1}$ and $p_{\text{astro}} \geq 0.9$	$SNR \geq 8$
GW150914	✓	✓
GW151012	✓	✓
GW151226	✓	✓
GW170104	✓	✓
GW170608	✓	✓
GW170729	✓	✓
GW170809	✓	✓
GW170814	✓	✓
GW170818	✓	✓
GW170823	✓	✓
GW190408_181802	✓	✓
GW190412	✓	✓
GW190413_052954		✓
GW190413_134308	✓	✓
GW190421_213856	✓	✓
GW190424_180648		✓
GW190426_190642		✓
GW190503_185404	✓	✓
GW190512_180714	✓	✓
GW190513_205428	✓	✓
GW190517_055101	✓	✓
GW190519_153544	✓	✓
GW190521_074359	✓	✓
GW190521	✓	✓
GW190527_092055		✓
GW190602_175927	✓	✓
GW190620_030421	✓	✓
GW190630_185205	✓	✓
GW190701_203306	✓	✓
GW190706_222641	✓	✓
GW190707_093326	✓	✓
GW190708_232457	✓	✓
GW190719_215514		✓
GW190720_000836	✓	✓
GW190725_174728		✓
GW190727_060333	✓	✓
GW190728_064510	✓	✓
GW190731_140936		✓
GW190803_022701	✓	✓
GW190805_211137		✓
GW190814	✓	✓
GW190828_063405	✓	✓
GW190828_065509	✓	✓
GW190910_112807	✓	✓
GW190915_235702	✓	✓
GW190916_200658		✓
GW190917_114630		✓
GW190924_021846	✓	✓
GW190925_232845	✓	✓
GW190926_050336		✓
GW190929_012149		✓
GW190930_133541	✓	✓
GW191103_012549		✓
GW191105_143521	✓	✓
GW191109_010717	✓	✓
GW191126_115259		✓
GW191127_050227		✓
GW191129_134029	✓	✓
GW191204_110529		✓
GW191204_171526	✓	✓
GW191215_223052	✓	✓
GW191216_213338	✓	✓
GW191222_033537	✓	✓
GW191230_180458	✓	✓
GW200112_155838	✓	✓
GW200128_022011	✓	✓
GW200129_065458	✓	✓
GW200202_154313	✓	✓
GW200208_130117	✓	✓
GW200209_085452	✓	✓
GW200210_092254		✓
GW200216_220804		✓
GW200219_094415	✓	✓
GW200220_124850		✓
GW200224_222234	✓	✓
GW200225_060421	✓	✓
GW200302_015811	✓	✓
GW200311_115853	✓	✓
GW200316_215756	✓	✓

Table B1. List of GW events included in the analysis for each selection criteria.

Model	CHE	Remnant mass prescription	$f_{WR}$	$\gamma$ prescription	$f_{\gamma}$	$\alpha_{CE}$	$M_{\alpha}$ prescription	$f_{MT}$	$\log_{10} B(\lambda/H)$	$\log_{10} B(\lambda/H)$ (SNR $\geq 8$ )	$\log_{10} B(\lambda, N_{\lambda}/H)$	$\log_{10} B(\lambda, N_{\lambda}/H)$ (SNR $\geq 8$ )
Rauf et al (2023) (a)	NONE	FRYER2012	1	ISOTROPIC	0	1	THERMAL	1	-10.1 $\pm$ 1.7	-10.9 $\pm$ 1.6	-66.8 $\pm$ 6.1	-57.3 $\pm$ 6.8
<b>Rauf et al (2023) (b)</b>	<b>NONE</b>	<b>MULLERMANDEL</b>	<b>1</b>	<b>ISOTROPIC</b>	<b>0</b>	<b>1</b>	<b>THERMAL</b>	<b>1</b>	<b>-3 <math>\pm</math> 2.2</b>	<b>-0.4 <math>\pm</math> 2.5</b>	<b>0 <math>\pm</math> 2.7</b>	<b>-0.3 <math>\pm</math> 4.6</b>
Rauf et al (2023) (c)	NONE	MULLERMANDEL	0.2	ISOTROPIC	0	1	THERMAL	1	-30.8 $\pm$ 3.4	-46.7 $\pm$ 4.4	-41.8 $\pm$ 5.9	-53.2 $\pm$ 7.1
Rauf et al (2023) (d)	NONE	SCHNEIDER2020	1	ISOTROPIC	0	1	THERMAL	1	-39.3 $\pm$ 8.6	-50.5 $\pm$ 10.6	-39.5 $\pm$ 8.8	-57.8 $\pm$ 11.4
1	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	0.1	FIXED	0.1	-37.7 $\pm$ 3.8	-61.5 $\pm$ 2.8	-56.6 $\pm$ 4.9	-59.9 $\pm$ 5.5
2	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	0.1	FIXED	0.2	-22.7 $\pm$ 3.6	-37 $\pm$ 2.6	-45.5 $\pm$ 7.2	-53.1 $\pm$ 7.6
3	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	0.1	FIXED	0.5	-26.8 $\pm$ 2.7	-44.1 $\pm$ 1.9	-195.6 $\pm$ 15.2	-196.9 $\pm$ 15.4
4	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	0.1	FIXED	1	-25.7 $\pm$ 2.2	-43.3 $\pm$ 1.4	-325.1 $\pm$ 18.8	-322.9 $\pm$ 19.1
5	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	1	FIXED	0.1	-62.3 $\pm$ 1.3	-83.3 $\pm$ 0.7	-114.4 $\pm$ 3.7	-125.5 $\pm$ 5
6	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	1	FIXED	0.2	-34 $\pm$ 1.3	-52.7 $\pm$ 0.7	-126.2 $\pm$ 4.7	-132.1 $\pm$ 5.8
7	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	1	FIXED	0.5	-27.7 $\pm$ 1.5	-37 $\pm$ 0.8	-257.3 $\pm$ 10	-248.6 $\pm$ 10.6
8	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	1	FIXED	1	-24.3 $\pm$ 1.8	-34.3 $\pm$ 1.1	-363 $\pm$ 17.3	-352.3 $\pm$ 17.6
9	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	10	FIXED	0.1	-43.9 $\pm$ 6.1	-66.1 $\pm$ 4.8	-42.3 $\pm$ 7.6	-64.5 $\pm$ 7.5
10	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	10	FIXED	0.2	-24.5 $\pm$ 4.9	-41.8 $\pm$ 3.7	-55.2 $\pm$ 10.4	-64.8 $\pm$ 10.6
11	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	10	FIXED	0.5	-23.9 $\pm$ 2.9	-41 $\pm$ 2	-210.2 $\pm$ 17.3	-210.5 $\pm$ 17.5
12	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	10	FIXED	1	-24.9 $\pm$ 2.1	-33 $\pm$ 1.3	-335.5 $\pm$ 18.4	-323.6 $\pm$ 18.7
13	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	100	FIXED	0.1	-34.9 $\pm$ 8.4	-53.1 $\pm$ 6.8	-38.2 $\pm$ 8.7	-65.8 $\pm$ 8.1
14	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	100	FIXED	0.2	-16.6 $\pm$ 6.1	-34.4 $\pm$ 4.9	-29 $\pm$ 10.1	-41.9 $\pm$ 10.1
15	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	100	FIXED	0.5	-24.7 $\pm$ 3.2	-35.2 $\pm$ 2.3	-194.1 $\pm$ 17.5	-188.3 $\pm$ 17.7
16	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	100	FIXED	1	-25.8 $\pm$ 2.5	-38 $\pm$ 1.7	-325.4 $\pm$ 21.7	-317.7 $\pm$ 21.9
17	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	0.1	FIXED	0.1	-8.8 $\pm$ 2.5	-15 $\pm$ 1.6	-166.8 $\pm$ 13.3	-157.2 $\pm$ 13.6
18	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	0.1	FIXED	0.2	-12.2 $\pm$ 1.8	-19.7 $\pm$ 1.1	-299.1 $\pm$ 15.1	-287.1 $\pm$ 15.5
19	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	0.1	FIXED	0.5	-17.4 $\pm$ 1.5	-24.9 $\pm$ 0.8	-566.4 $\pm$ 19.6	-550 $\pm$ 19.9
20	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	0.1	FIXED	1	-11.6 $\pm$ 1.5	-16.5 $\pm$ 0.8	-505.8 $\pm$ 19	-487.5 $\pm$ 19.3
21	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	1	FIXED	0.1	-15.7 $\pm$ 1.3	-23.3 $\pm$ 0.7	-256.2 $\pm$ 7.8	-245.5 $\pm$ 8.5
22	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	1	FIXED	0.2	-8.6 $\pm$ 1.2	-15.7 $\pm$ 0.8	-374 $\pm$ 9.1	-359.9 $\pm$ 9.8
23	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	1	FIXED	0.5	-10.1 $\pm$ 1.2	-16.1 $\pm$ 0.8	-564.2 $\pm$ 13.4	-546.2 $\pm$ 13.9
24	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	1	FIXED	1	-7.4 $\pm$ 1.4	-11.2 $\pm$ 0.8	-440.3 $\pm$ 15.6	-421.9 $\pm$ 16
25	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	10	FIXED	0.1	-17.2 $\pm$ 2.4	-26.4 $\pm$ 1.6	-205.3 $\pm$ 14.5	-197.6 $\pm$ 14.9
26	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	10	FIXED	0.2	-26.4 $\pm$ 1.8	-40.5 $\pm$ 1.1	-338.9 $\pm$ 16.4	-332.9 $\pm$ 16.7
27	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	10	FIXED	0.5	-22.5 $\pm$ 1.4	-30.9 $\pm$ 0.8	-594.5 $\pm$ 20	-578.6 $\pm$ 20.3
28	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	10	FIXED	1	-14.9 $\pm$ 1.5	-19.6 $\pm$ 0.8	-470.7 $\pm$ 17.7	-452.7 $\pm$ 18.1
29	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	100	FIXED	0.1	-13.2 $\pm$ 2.9	-31.7 $\pm$ 2	-162.7 $\pm$ 14.8	-165.7 $\pm$ 15.1
30	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	100	FIXED	0.2	-24.4 $\pm$ 2	-35.8 $\pm$ 1.2	-319.4 $\pm$ 17.4	-311.1 $\pm$ 17.7
31	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	100	FIXED	0.5	-16 $\pm$ 1.6	-23.6 $\pm$ 0.9	-506.4 $\pm$ 20.6	-490.8 $\pm$ 20.9
32	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	100	FIXED	1	-14.4 $\pm$ 1.8	-21.3 $\pm$ 1.1	-443.8 $\pm$ 20.8	-428.2 $\pm$ 21.1
33	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	0.1	FIXED	0.1	-10.9 $\pm$ 1.2	-16.2 $\pm$ 0.7	-499.5 $\pm$ 13	-481.7 $\pm$ 13.4
34	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	0.1	FIXED	0.2	-6.9 $\pm$ 1.2	-11.2 $\pm$ 0.7	-442 $\pm$ 12.1	-424 $\pm$ 12.6
35	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	0.1	FIXED	0.5	-3.5 $\pm$ 1.2	-7.1 $\pm$ 0.8	-569.3 $\pm$ 13.7	-548.8 $\pm$ 14.1
36	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	0.1	FIXED	1	-3.2 $\pm$ 1.7	-6 $\pm$ 0.9	-339.7 $\pm$ 15.6	-321.9 $\pm$ 15.9
37	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	1	FIXED	0.1	-7.3 $\pm$ 1.2	-11.7 $\pm$ 0.9	-522.6 $\pm$ 9.4	-503.5 $\pm$ 9.4
38	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	1	FIXED	0.2	-8.9 $\pm$ 1.2	-13.4 $\pm$ 0.8	-476.7 $\pm$ 9.2	-458.3 $\pm$ 9.9
39	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	1	FIXED	0.5	-3.1 $\pm$ 1.2	-6 $\pm$ 0.9	-553 $\pm$ 9.9	-532 $\pm$ 10.5
40	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	1	FIXED	1	-2.4 $\pm$ 1.5	-3.3 $\pm$ 0.9	-378.3 $\pm$ 15.4	-358 $\pm$ 15.8
41	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	10	FIXED	0.1	-8.6 $\pm$ 1.2	-12.9 $\pm$ 0.8	-470.3 $\pm$ 10.9	-451.8 $\pm$ 11.4
42	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	10	FIXED	0.2	-8.8 $\pm$ 1.2	-13.7 $\pm$ 0.8	-435.8 $\pm$ 9.4	-418.4 $\pm$ 10
43	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	10	FIXED	0.5	-5.6 $\pm$ 1.2	-8.5 $\pm$ 0.8	-526.7 $\pm$ 12.2	-506 $\pm$ 12.7
44	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	10	FIXED	1	0 $\pm$ 1.6	0 $\pm$ 0.9	-331.9 $\pm$ 14.2	-311.5 $\pm$ 14.6
45	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	100	FIXED	0.1	13.6 $\pm$ 1.2	-20.6 $\pm$ 0.7	-466.2 $\pm$ 12.7	-450.6 $\pm$ 13.2
46	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	100	FIXED	0.2	-8.9 $\pm$ 1.3	-13.2 $\pm$ 0.7	-408.9 $\pm$ 11.9	-391.5 $\pm$ 12.4
47	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	100	FIXED	0.5	-3.7 $\pm$ 1.3	-6.4 $\pm$ 0.7	-500.9 $\pm$ 15.1	-480.4 $\pm$ 15.5
48	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	100	FIXED	1	-7.8 $\pm$ 2	-11.6 $\pm$ 1.2	-305.2 $\pm$ 17	-289.2 $\pm$ 17.3
49	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	0.1	FIXED	0.1	-10.2 $\pm$ 3.2	-13.4 $\pm$ 2.3	-122 $\pm$ 5.7	-141.4 $\pm$ 6.3
50	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	0.1	FIXED	0.2	-109.7 $\pm$ 3.1	-134.1 $\pm$ 2.2	-126.6 $\pm$ 5.9	-145.2 $\pm$ 6.5
51	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	0.1	FIXED	0.5	-23.2 $\pm$ 3	-41 $\pm$ 2.1	-75.9 $\pm$ 8.4	-83.7 $\pm$ 8.9
52	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	0.1	FIXED	1	-2 $\pm$ 2.8	-5.1 $\pm$ 2	-78.1 $\pm$ 9.6	-69.5 $\pm$ 10
53	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	1	FIXED	0.1	-123.3 $\pm$ 2.3	-151.5 $\pm$ 1.5	-152.2 $\pm$ 5.2	-172.9 $\pm$ 6
54	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	1	FIXED	0.2	-133.8 $\pm$ 2.3	-162.7 $\pm$ 1.5	-174.6 $\pm$ 5.9	-194.7 $\pm$ 6.7
55	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	1	FIXED	0.5	-40.8 $\pm$ 2.3	-61 $\pm$ 1.4	-93.2 $\pm$ 6.5	-103.4 $\pm$ 7.2
56	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	1	FIXED	1	-1.3 $\pm$ 2.4	-1 $\pm$ 1.6	-73.9 $\pm$ 8.1	-62.1 $\pm$ 8.7
57	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	10	FIXED	0.1	-119.8 $\pm$ 3.7	-148.9 $\pm$ 2.7	-119 $\pm$ 5	-147.6 $\pm$ 5.6
58	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	10	FIXED	0.2	-126.1 $\pm$ 3.6	-155.1 $\pm$ 2.6	-124.6 $\pm$ 4.7	-153.5 $\pm$ 5.4
59	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	10	FIXED	0.5	-42.5 $\pm$ 3.4	-62.9 $\pm$ 2.4	-53.1 $\pm$ 5.7	-69 $\pm$ 6.3
60	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	10	FIXED	1	-2.4 $\pm$ 3.2	-1.2 $\pm$ 2.3	-45.9 $\pm$ 8.2	-35.5 $\pm$ 8.7
61	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	100	FIXED	0.1	-124.4 $\pm$ 5	-154.3 $\pm$ 3.8	-124.5 $\pm$ 5.4	-161.8 $\pm$ 5.7
62	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	100	FIXED	0.2	-99.8 $\pm$ 4.5	-123.9 $\pm$ 3.4	-101 $\pm$ 4.9	-133.2 $\pm$ 5.3
63	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	100	FIXED	0.5	-19 $\pm$ 4.7	-34.6 $\pm$ 3.6	-16.1 $\pm$ 5.7	-33.3 $\pm$ 6
64	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	100	FIXED	1	-16.4 $\pm$ 5.3	-21.6 $\pm$ 4.2	-25 $\pm$ 8.4	-26.2 $\pm$ 8.5
65	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	0.1	FIXED	0.1	-112.6 $\pm$ 3.2	-137.3 $\pm$ 2.2	-163.4 $\pm$ 8.7	-178.3 $\pm$ 9.1
66	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	0.1	FIXED	0.2	-141.1 $\pm$ 3.6	-169 $\pm$ 2.6	-181.7 $\pm$ 8.8	-200.8 $\pm$ 9.2
67	NONE	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	0.1	FIXED	0.5	-121.5 $\pm$ 3.3	-148.7 $\pm$ 2.4	-166.5 $\pm$ 8.6	

Model	CHE	Remnant mass prescription	$f_{WR}$	$\gamma$ prescription	$f_{\gamma}$	$\alpha_{CE}$	$M_a$ prescription	$f_{MT}$	$\log_{10} \mathcal{B}(\lambda H)$	$\log_{10} \mathcal{B}(\lambda H)$ (SNR $\geq 8$ )	$\log_{10} \mathcal{B}(\lambda, N_{\lambda} H)$	$\log_{10} \mathcal{B}(\lambda, N_{\lambda} H)$ (SNR $\geq 8$ )
101	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	1	FIXED	0.1	-23.9 ± 1.2	-28 ± 0.7	-671.1 ± 16.5	-650.1 ± 16.9
102	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	1	FIXED	0.2	-27.6 ± 1.2	-32.5 ± 0.8	-918.5 ± 17.6	-895.9 ± 18
103	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	1	FIXED	0.5	-11.9 ± 1.2	-14.7 ± 0.8	-936.3 ± 17.7	-911.4 ± 18
104	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	1	FIXED	1	-11.4 ± 1.3	-14.4 ± 0.7	-941.1 ± 25.8	-916.3 ± 26
105	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	10	FIXED	0.1	-21.2 ± 2	-26.2 ± 1.2	-630.8 ± 31.4	-611 ± 31.5
106	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	10	FIXED	0.2	-22.6 ± 1.6	-27.9 ± 0.9	-824.5 ± 30.1	-803.2 ± 30.3
107	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	10	FIXED	0.5	-29.9 ± 1.4	-35 ± 0.8	-1013.1 ± 29.5	-989.9 ± 29.7
108	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	10	FIXED	1	-18.7 ± 1.3	-21.3 ± 0.8	-917.5 ± 26.3	-892.5 ± 26.5
109	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	100	FIXED	0.1	-32.6 ± 2.2	-39.1 ± 1.4	-577.3 ± 31.6	-560 ± 31.7
110	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	100	FIXED	0.2	-24.3 ± 1.7	-30 ± 1	-803.3 ± 31.8	-782.6 ± 32
111	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	100	FIXED	0.5	-30.6 ± 1.4	-37 ± 0.8	-1006.4 ± 30.8	-984.6 ± 31
112	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	100	FIXED	1	-27 ± 1.5	-31.8 ± 0.9	-876.4 ± 30.9	-854.1 ± 31.1
113	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	0.1	FIXED	0.1	-12.8 ± 1.2	-15 ± 0.8	-944.8 ± 17.6	-919.3 ± 17.9
114	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	0.1	FIXED	0.2	-12.9 ± 1.2	-15.7 ± 0.8	-962.4 ± 17.3	-937.2 ± 17.7
115	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	0.1	FIXED	0.5	-7.6 ± 1.2	-7.8 ± 0.8	-977.8 ± 17.5	-949.9 ± 17.8
116	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	0.1	FIXED	1	-24.2 ± 1.5	-29.1 ± 0.8	-752.7 ± 25.5	-731.6 ± 25.8
117	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	1	FIXED	0.1	-22 ± 1.2	-25.4 ± 1	-969.7 ± 11.3	-945.2 ± 11.9
118	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	1	FIXED	0.2	-21.7 ± 1.2	-26 ± 0.9	-1105 ± 13.8	-1080.4 ± 14.3
119	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	1	FIXED	0.5	-10.5 ± 1.2	-12.1 ± 0.9	-1058.6 ± 14.6	-1031.6 ± 15
120	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	1	FIXED	1	-9.6 ± 1.3	-10.9 ± 0.7	-777 ± 21.7	-751.9 ± 21.9
121	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	10	FIXED	0.1	-16.4 ± 1.2	-19.9 ± 0.9	-936 ± 13.7	-911.7 ± 14.1
122	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	10	FIXED	0.2	-13.7 ± 1.2	-16.3 ± 0.9	-914.3 ± 14.5	-889.3 ± 14.9
123	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	10	FIXED	0.5	-11.1 ± 1.2	-12.5 ± 0.9	-998.7 ± 15.8	-971.9 ± 16.1
124	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	10	FIXED	1	-13.5 ± 1.4	-15.7 ± 0.8	-780.6 ± 24	-756.4 ± 24.2
125	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	100	FIXED	0.1	-14.1 ± 1.2	-16.7 ± 0.8	-840.1 ± 18.5	-815.8 ± 18.8
126	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	100	FIXED	0.2	-22.5 ± 1.2	-27.1 ± 0.8	-848.5 ± 18.6	-826.1 ± 18.9
127	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	100	FIXED	0.5	-9.8 ± 1.2	-10.2 ± 0.8	-934.8 ± 21.3	-907.5 ± 21.6
128	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	100	FIXED	1	-20.8 ± 1.7	-25.8 ± 1	-797.8 ± 31.9	-776.3 ± 32.1
129	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	0.1	FIXED	0.1	-27.7 ± 2.4	-32 ± 1.6	-452 ± 27.4	-434.1 ± 27.5
130	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	0.1	FIXED	0.2	-29.6 ± 2.5	-36 ± 1.6	-427.1 ± 26.5	-411.8 ± 26.7
131	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	0.1	FIXED	0.5	-32.2 ± 2.5	-39.8 ± 1.6	-556.2 ± 33.3	-540.2 ± 33.5
132	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	0.1	FIXED	1	-13.9 ± 2.1	-17.2 ± 1.3	-554.2 ± 28.9	-533.7 ± 29.1
133	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	1	FIXED	0.1	-26.4 ± 1.9	-30.6 ± 1.1	-369.1 ± 18.3	-352.6 ± 18.6
134	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	1	FIXED	0.2	-26.6 ± 1.9	-32.9 ± 1.1	-435 ± 20.8	-419.4 ± 21.1
135	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	1	FIXED	0.5	-16.6 ± 1.9	-19.6 ± 1.1	-487.8 ± 23.2	-468 ± 23.5
136	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	1	FIXED	1	-22.8 ± 2	-26.1 ± 1.3	-517.7 ± 26.7	-497.8 ± 26.9
137	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	10	FIXED	0.1	-37.7 ± 2.8	-45 ± 1.9	-424.6 ± 29.1	-410.4 ± 29.2
138	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	10	FIXED	0.2	-37.7 ± 2.6	-45.1 ± 1.8	-425 ± 27.5	-410.8 ± 27.6
139	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	10	FIXED	0.5	-25.5 ± 2.5	-31.1 ± 1.7	-482.9 ± 30.3	-465.9 ± 30.4
140	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	10	FIXED	1	-21.1 ± 2.5	-24.2 ± 1.7	-410.6 ± 26.8	-392.1 ± 26.9
141	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	100	FIXED	0.1	-40.3 ± 3	-47.6 ± 2.1	-429.5 ± 31.6	-415.1 ± 31.8
142	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	100	FIXED	0.2	-49.1 ± 3.1	-58 ± 2.2	-499.1 ± 36.1	-485.5 ± 36.2
143	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	100	FIXED	0.5	-43.2 ± 3.2	-52.7 ± 2.3	-470.7 ± 35.8	-458 ± 35.9
144	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	100	FIXED	1	-39.7 ± 3.2	-47.9 ± 2.3	-503.2 ± 38.6	-488.6 ± 38.7
145	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	0.1	FIXED	0.1	-25.9 ± 2.6	-33.9 ± 1.7	-491.5 ± 31.6	-476.7 ± 31.8
146	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	0.1	FIXED	0.2	-25.7 ± 2.6	-32.7 ± 1.7	-497 ± 32.3	-481.2 ± 32.4
147	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	0.1	FIXED	0.5	-33.3 ± 2.3	-41.6 ± 1.5	-557 ± 31.7	-541.7 ± 31.9
148	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	0.1	FIXED	1	-18.2 ± 2.3	-24.2 ± 1.5	-510.1 ± 29.8	-492.9 ± 30
149	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	1	FIXED	0.1	-32.9 ± 2.1	-40 ± 1.3	-513.2 ± 26.8	-497.3 ± 27
150	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	1	FIXED	0.2	-31.5 ± 2.1	-38.8 ± 1.3	-504.2 ± 26.2	-488.7 ± 26.4
151	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	1	FIXED	0.5	-31.6 ± 2	-37.4 ± 1.2	-558.6 ± 27.6	-540.7 ± 27.8
152	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	1	FIXED	1	-29.9 ± 2	-36.9 ± 1.2	-525.7 ± 26.3	-509.4 ± 26.5
153	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	10	FIXED	0.1	-43.7 ± 3.1	-53.1 ± 2.2	-500.7 ± 37	-487.5 ± 37.1
154	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	10	FIXED	0.2	-36 ± 3.1	-44.7 ± 2.2	-462.4 ± 34.4	-448.9 ± 34.5
155	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	10	FIXED	0.5	-34.7 ± 3.1	-43 ± 2.2	-401.5 ± 31	-388.6 ± 31.2
156	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	10	FIXED	1	-38.6 ± 2.6	-47.4 ± 1.7	-532.7 ± 33.4	-518.3 ± 33.5
157	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	100	FIXED	0.1	-63.8 ± 3.3	-84.5 ± 2.4	-542.5 ± 40.5	-540.2 ± 40.6
158	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	100	FIXED	0.2	-60.3 ± 3.5	-77.2 ± 2.5	-443.1 ± 35.2	-438.5 ± 35.3
159	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	100	FIXED	0.5	-38.4 ± 3.3	-50.2 ± 2.4	-487.3 ± 38.4	-476.6 ± 38.5
160	PESSIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	100	FIXED	1	-52.7 ± 3.4	-66.9 ± 2.5	-549.6 ± 43.3	-540.5 ± 43.4
161	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	0.1	FIXED	0.1	-56.6 ± 2.4	-67.9 ± 1.6	-853.3 ± 47	-837.9 ± 47.1
162	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	0.1	FIXED	0.2	-52.2 ± 2.4	-63 ± 1.6	-990.4 ± 54.8	-973.4 ± 54.9
163	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	0.1	FIXED	0.5	-30.8 ± 2	-37.8 ± 1.2	-970.6 ± 45.7	-949.8 ± 45.8
164	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	0.1	FIXED	1	-32.8 ± 1.7	-40.6 ± 1	-1043.1 ± 40.2	-1022.5 ± 40.3
165	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	1	FIXED	0.1	-30.6 ± 1.2	-35.4 ± 0.8	-860.2 ± 17.3	-838.1 ± 17.7
166	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	1	FIXED	0.2	-38.7 ± 1.2	-45.4 ± 0.8	-906.4 ± 17.6	-885.9 ± 18
167	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	1	FIXED	0.5	-30.3 ± 1.3	-38 ± 0.7	-1111.1 ± 29.2	-1089.8 ± 29.4
168	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	1	FIXED	1	-39.2 ± 1.5	-47.7 ± 0.8	-1139.6 ± 38.5	-1118.9 ± 38.6
169	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	10	FIXED	0.1	-55.5 ± 3	-65.8 ± 2.1	-925.8 ± 62.2	-908.8 ± 62.2
170	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	10	FIXED	0.2	-50.8 ± 2.8	-60.9 ± 1.9	-999 ± 63	-981.2 ± 63.1
171	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	10	FIXED	0.5	-27.8 ± 2	-34.7 ± 1.2	-1146.8 ± 52.9	-1124.4 ± 53
172	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	10	FIXED	1	-29.2 ± 1.6	-35.2 ± 0.9	-1118.6 ± 40.8	-1095.6 ± 40.9
173	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	100	FIXED	0.1	-63.9 ± 3.3	-76.6 ± 2.4	-680.9 ± 50.8	-668.7 ± 50.9
174	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	100	FIXED	0.2	-65.9 ± 2.9	-82 ± 2	-999.6 ± 64.6	-987.9 ± 64.7
175	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	100	FIXED	0.5	-33.9 ± 2.4	-43.2 ± 1.5	-872 ± 48.4	-854.3 ± 48.5
176	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0	100	FIXED	1	-42.3 ± 1.9	-53.4 ± 1.2	-1169.6 ± 52.1	-1151.5 ± 52.2
177	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	0.1	FIXED	0.1	-44.7 ± 1.7	-53.9 ± 1	-1033.7 ± 41.5	-1014.6 ± 41.7
178	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	0.1	FIXED	0.2	-39.6 ± 1.5	-47.4 ± 0.8	-1170.3 ± 37.9	-1148.8 ± 38.1
179	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	0.1	FIXED	0.5	-30.5 ± 1.2	-36.7 ± 0.7	-1471.8 ± 35.2	-1446.9 ± 35.4
180	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	0.1	FIXED	1	-37.6 ± 1.3	-43.8 ± 0.7	-1286 ± 32.7	-1262.2 ± 32.7
181	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	1	FIXED	0.1	-27.7 ± 1.2	-32.5 ± 0.8	-1055.2 ± 22.5	-1031.5 ± 22.8
182	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	1	FIXED	0.2	-26.3 ± 1.2	-32.2 ± 0.9	-1305.9 ± 20	-1281.6 ± 20.3
183	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	1	FIXED	0.5	-28.7 ± 1.2	-33.7 ± 0.8	-1413.4 ± 23.6	-1387.6 ± 23.9
184	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	1	FIXED	1	-13.3 ± 1.3	-16.5 ± 0.7	-1214.8 ± 31.7	-1188.3 ± 31.9
185	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	10	FIXED	0.1	-37 ± 1.9	-44.4 ± 1.1	-1106.3 ± 47.6	-1084.9 ± 47.8
186	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	10	FIXED	0.2	-34.2 ± 1.5	-40.6 ± 0.8	-1144.1 ± 37.7	-1121.4 ± 37.8
187	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	10	FIXED	0.5	-28.2 ± 1.3	-28.2 ± 0.7	-1450.4 ± 36	-1424.4 ± 36.2
188	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	10	FIXED	1	-19.6 ± 1.3	-23.7 ± 0.7	-1276.4 ± 33.5	-1250.4 ± 33.7
189	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.25	100	FIXED					

Model	CHE	Remnant mass prescription	$f_{WR}$	$\gamma$ prescription	$f_{\gamma}$	$\alpha_{CE}$	$M_{\alpha}$ prescription	$f_{MT}$	$\log_{10} \mathcal{B}(\lambda H)$	$\log_{10} \mathcal{B}(\lambda H)$ (SNR $\geq 8$ )	$\log_{10} \mathcal{B}(\lambda, N_{\lambda} H)$	$\log_{10} \mathcal{B}(\lambda, N_{\lambda} H)$ (SNR $\geq 8$ )
201	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	10	FIXED	0.1	-17.5 ± 1.2	-20.4 ± 0.9	-1226.5 ± 18.2	-1199.6 ± 18.5
202	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	10	FIXED	0.2	-16.4 ± 1.2	-20.2 ± 0.9	-1404.4 ± 18.6	-1377.4 ± 18.9
203	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	10	FIXED	0.5	-14.6 ± 1.2	-17.7 ± 0.9	-1495.4 ± 21.6	-1467.2 ± 21.8
204	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	10	FIXED	1	-18.7 ± 1.3	-22.2 ± 0.8	-1176.7 ± 32.9	-1150.8 ± 33.1
205	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	100	FIXED	0.1	-14.7 ± 1.2	-17.6 ± 0.8	-1269.5 ± 22	-1242.4 ± 22.3
206	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	100	FIXED	0.2	-15.2 ± 1.2	-18.5 ± 0.8	-1312.8 ± 24.5	-1285.8 ± 24.8
207	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	100	FIXED	0.5	-14.2 ± 1.2	-16.9 ± 0.8	-1335.9 ± 27.1	-1306.2 ± 27.3
208	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.5	100	FIXED	1	-23.2 ± 1.6	-29.2 ± 0.9	-1156.9 ± 41.5	-1133.6 ± 41.6
209	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	0.1	FIXED	0.1	-24.8 ± 2.2	-31 ± 1.4	-979.3 ± 50.1	-957.5 ± 50.2
210	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	0.1	FIXED	0.2	-33.3 ± 2.2	-39.7 ± 1.4	-846 ± 44	-825.6 ± 44.1
211	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	0.1	FIXED	0.5	-23.5 ± 2.3	-28.8 ± 1.4	-823.9 ± 44.5	-802.5 ± 44.6
212	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	0.1	FIXED	1	-20.1 ± 2.1	-25.5 ± 1.3	-875.1 ± 44.1	-853.3 ± 44.2
213	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	1	FIXED	0.1	-21 ± 1.8	-25.1 ± 1.1	-743.4 ± 32.7	-721.5 ± 32.9
214	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	1	FIXED	0.2	-31.3 ± 1.7	-36.1 ± 1	-1014.4 ± 39.4	-991 ± 39.5
215	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	1	FIXED	0.5	-23.3 ± 1.8	-27.5 ± 1.1	-956.8 ± 40.8	-933.1 ± 40.9
216	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	1	FIXED	1	-22.5 ± 1.9	-27.1 ± 1.1	-1022.9 ± 44.9	-999.2 ± 45
217	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	10	FIXED	0.1	-29.7 ± 2.4	-35.7 ± 1.6	-851.6 ± 49.1	-830.7 ± 49.2
218	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	10	FIXED	0.2	-28 ± 2.4	-33.1 ± 1.5	-771.2 ± 43.8	-750.3 ± 44
219	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	10	FIXED	0.5	-23.1 ± 2.2	-26.6 ± 1.4	-771.1 ± 41	-748.4 ± 41.1
220	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	10	FIXED	1	-26.9 ± 2.1	-31.8 ± 1.3	-910.7 ± 46.3	-888.1 ± 46.4
221	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	100	FIXED	0.1	-45.1 ± 2.9	-68.1 ± 2	-831.5 ± 55.4	-827.9 ± 55.5
222	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	100	FIXED	0.2	-46.6 ± 2.8	-56.1 ± 1.9	-750.8 ± 48.4	-734.6 ± 48.5
223	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	100	FIXED	0.5	-30.1 ± 2.7	-36.6 ± 1.8	-843.9 ± 53.6	-823.3 ± 53.7
224	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	0.75	100	FIXED	1	-61.6 ± 2.9	-72.5 ± 2	-883.8 ± 57.8	-867.9 ± 57.9
225	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	0.1	FIXED	0.1	-37.3 ± 2.3	-45.3 ± 1.5	-936.9 ± 50	-917.4 ± 50
226	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	0.1	FIXED	0.2	-26.9 ± 2.3	-34.3 ± 1.5	-756.1 ± 41.9	-737.5 ± 42
227	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	0.1	FIXED	0.5	-37.3 ± 2.3	-46.9 ± 1.5	-868.2 ± 47.7	-850.8 ± 47.8
228	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	0.1	FIXED	1	-19.5 ± 2	-25.2 ± 1.2	-865.2 ± 41.8	-843.8 ± 41.9
229	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	1	FIXED	0.1	-31.5 ± 1.9	-37.8 ± 1.1	-859 ± 38.5	-838.4 ± 38.6
230	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	1	FIXED	0.2	-35.3 ± 1.9	-43.5 ± 1.1	-1031.6 ± 45.4	-1011.5 ± 45.5
231	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	1	FIXED	0.5	-32.8 ± 1.9	-40.9 ± 1.1	-863 ± 38.2	-844.1 ± 38.3
232	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	1	FIXED	1	-34.5 ± 1.9	-41.1 ± 1.1	-974 ± 42.8	-952.7 ± 42.9
233	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	10	FIXED	0.1	-55.5 ± 2.9	-68.2 ± 2	-766.6 ± 50.5	-753.5 ± 50.6
234	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	10	FIXED	0.2	-56.7 ± 2.6	-68.8 ± 1.7	-967.7 ± 56.8	-952 ± 56.9
235	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	10	FIXED	0.5	-43.7 ± 2.7	-53.9 ± 1.8	-844.4 ± 52.2	-827.9 ± 52.3
236	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	10	FIXED	1	-25 ± 2.4	-31 ± 1.5	-1012.4 ± 56.2	-990.2 ± 56.3
237	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	100	FIXED	0.1	-62.3 ± 3	-80 ± 2.1	-886.3 ± 59.5	-877.1 ± 59.5
238	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	100	FIXED	0.2	-52.9 ± 3.1	-65.6 ± 2.2	-820.6 ± 57.8	-807 ± 57.9
239	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	100	FIXED	0.5	-59.5 ± 3.1	-75.1 ± 2.4	-844.3 ± 58	-833.4 ± 58
240	OPTIMISTIC	MULLERMANDEL	0.2	MACLEOD_LINEAR	1	100	FIXED	1	-44.2 ± 3	-56.1 ± 2.1	-741.1 ± 50.8	-727.4 ± 50.8
241	NONE	MULLERMANDEL	0.1	ISOTROPIC	0	1	THERMAL	0.5	-27.6 ± 1.4	-42.4 ± 0.8	-347.4 ± 12.2	-341.7 ± 12.6
242	NONE	MULLERMANDEL	0.5	ISOTROPIC	0	1	THERMAL	0.5	-18 ± 1.5	-31.7 ± 0.9	-246.3 ± 10.4	-242 ± 10.9
243	NONE	MULLERMANDEL	1	ISOTROPIC	0	1	THERMAL	0.5	-11.7 ± 1.8	-18.8 ± 1	-143.3 ± 8.7	-132.7 ± 9.3
244	NONE	MULLERMANDEL	5	ISOTROPIC	0	1	THERMAL	0.5	-6.1 ± 2.8	-11.2 ± 1.9	-10.4 ± 4.4	-12.7 ± 5.3
245	NONE	MULLERMANDEL	10	ISOTROPIC	0	1	THERMAL	0.5	-13.5 ± 3.7	-15.3 ± 2.7	-10.4 ± 4.4	-15.5 ± 5.1
246	NONE	SCHNEIDER2020	0.1	ISOTROPIC	0	1	THERMAL	0.5	-18.3 ± 5.1	-48.7 ± 3.9	-46.5 ± 10.5	-69.5 ± 10.6
247	NONE	SCHNEIDER2020	0.5	ISOTROPIC	0	1	THERMAL	0.5	-28.3 ± 5.5	-61.8 ± 4.3	-42 ± 9.3	-70.5 ± 9.4
248	NONE	SCHNEIDER2020	1	ISOTROPIC	0	1	THERMAL	0.5	-12 ± 6.1	-39.5 ± 4.8	-11 ± 7.8	-38.1 ± 7.7
249	NONE	SCHNEIDER2020	5	ISOTROPIC	0	1	THERMAL	0.5	-6.6 ± 8.2	-33 ± 6.7	-15 ± 8.5	-53.4 ± 7.8
250	NONE	SCHNEIDER2020	10	ISOTROPIC	0	1	THERMAL	0.5	-14.2 ± 9	-41.6 ± 7.4	-38.8 ± 9.2	-84.8 ± 8.4
251	PESSIMISTIC	MULLERMANDEL	0.1	ISOTROPIC	0	1	THERMAL	0.5	-29.2 ± 1.3	-34.9 ± 0.7	-793.7 ± 21.1	-773.1 ± 21.4
252	PESSIMISTIC	MULLERMANDEL	0.5	ISOTROPIC	0	1	THERMAL	0.5	-30.6 ± 1.4	-34.9 ± 0.8	-636.8 ± 20	-616.4 ± 20.3
253	PESSIMISTIC	MULLERMANDEL	1	ISOTROPIC	0	1	THERMAL	0.5	-35.9 ± 1.6	-40.3 ± 0.9	-501.7 ± 18.7	-483.2 ± 19
254	PESSIMISTIC	MULLERMANDEL	5	ISOTROPIC	0	1	THERMAL	0.5	-19.6 ± 2.6	-23.7 ± 1.7	-77.7 ± 7.6	-71.3 ± 8.2
255	PESSIMISTIC	MULLERMANDEL	10	ISOTROPIC	0	1	THERMAL	0.5	-19.4 ± 3.5	-21.8 ± 2.5	-27.6 ± 5.7	-26.1 ± 6.3
256	PESSIMISTIC	SCHNEIDER2020	0.1	ISOTROPIC	0	1	THERMAL	0.5	-53.1 ± 3.2	-78.9 ± 2.3	-480.4 ± 36.1	-484 ± 36.2
257	PESSIMISTIC	SCHNEIDER2020	0.5	ISOTROPIC	0	1	THERMAL	0.5	-59 ± 3.5	-87.9 ± 2.5	-455.2 ± 36.3	-462.4 ± 36.4
258	PESSIMISTIC	SCHNEIDER2020	1	ISOTROPIC	0	1	THERMAL	0.5	-50.9 ± 3.9	-71 ± 2.9	-314.8 ± 28.9	-315.9 ± 29
259	PESSIMISTIC	SCHNEIDER2020	5	ISOTROPIC	0	1	THERMAL	0.5	-53.2 ± 6	-88.7 ± 4.8	-68.4 ± 10.4	-98.4 ± 10.4
260	PESSIMISTIC	SCHNEIDER2020	10	ISOTROPIC	0	1	THERMAL	0.5	-51.9 ± 7.2	-87.3 ± 5.7	-49.9 ± 8.8	-85.7 ± 8.5
261	OPTIMISTIC	MULLERMANDEL	0.1	ISOTROPIC	0	1	THERMAL	0.5	-33.9 ± 1.2	-41.6 ± 0.7	-1324.5 ± 30.8	-1301.8 ± 31
262	OPTIMISTIC	MULLERMANDEL	0.5	ISOTROPIC	0	1	THERMAL	0.5	-37.3 ± 1.4	-42.5 ± 0.8	-845 ± 25.4	-823.5 ± 25.7
263	OPTIMISTIC	MULLERMANDEL	1	ISOTROPIC	0	1	THERMAL	0.5	-35.6 ± 1.6	-41.1 ± 0.9	-577.5 ± 21.4	-559.1 ± 21.6
264	OPTIMISTIC	MULLERMANDEL	5	ISOTROPIC	0	1	THERMAL	0.5	-26.1 ± 2.6	-29.5 ± 1.7	-89.8 ± 8.1	-82.3 ± 8.6
265	OPTIMISTIC	MULLERMANDEL	10	ISOTROPIC	0	1	THERMAL	0.5	-34.1 ± 3.5	-41.1 ± 2.5	-44 ± 5.8	-46.6 ± 6.4
266	OPTIMISTIC	SCHNEIDER2020	0.1	ISOTROPIC	0	1	THERMAL	0.5	-79.7 ± 2.7	-120.1 ± 1.9	-1191.7 ± 71.8	-1202.9 ± 71.8
267	OPTIMISTIC	SCHNEIDER2020	0.5	ISOTROPIC	0	1	THERMAL	0.5	-59.9 ± 3.2	-96.4 ± 2.3	-553.9 ± 40.8	-567.3 ± 40.9
268	OPTIMISTIC	SCHNEIDER2020	1	ISOTROPIC	0	1	THERMAL	0.5	-66.5 ± 3.6	-98.8 ± 2.7	-419.8 ± 34.5	-431.3 ± 34.6
269	OPTIMISTIC	SCHNEIDER2020	5	ISOTROPIC	0	1	THERMAL	0.5	-62.8 ± 5.9	-97.9 ± 4.6	-81.2 ± 10.7	-110.3 ± 10.6
270	OPTIMISTIC	SCHNEIDER2020	10	ISOTROPIC	0	1	THERMAL	0.5	-45.9 ± 7.2	-81.4 ± 5.8	-44.8 ± 9	-79.9 ± 8.7
271	NONE	FRYER2012	0.1	ISOTROPIC	0	1	THERMAL	0.5	-14.2 ± 2.2	-19.4 ± 1.4	-209.6 ± 13.5	-197.7 ± 13.8
272	NONE	FRYER2012	0.5	ISOTROPIC	0	1	THERMAL	0.5	-11.7 ± 2.3	-13.5 ± 1.5	-150.5 ± 11.3	-137.3 ± 11.8
273	NONE	FRYER2012	1	ISOTROPIC	0	1	THERMAL	0.5	-4.4 ± 2.7	-6.7 ± 1.8	-86.4 ± 9.4	-76.5 ± 9.9
274	NONE	FRYER2012	5	ISOTROPIC	0	1	THERMAL	0.5	-1.9 ± 3.8	-1.5 ± 2.8	0 ± 5.6	0 ± 5.6
275	NONE	FRYER2012	10	ISOTROPIC	0	1	THERMAL	0.5	-3.8 ± 4.7	-2.5 ± 3.6	-2.9 ± 5.2	-8.1 ± 5.5
276	NONE	FRYER2022	0.1	ISOTROPIC	0	1	THERMAL	0.5	-15.9 ± 2.2	-20.6 ± 1.4	-210.8 ± 13.9	-198.5 ± 14.2
277	NONE	FRYER2022	0.5	ISOTROPIC	0	1	THERMAL	0.5	-14 ± 2.5	-19.5 ± 1.6	-144.7 ± 11.7	-135.5 ± 12.1
278	NONE	FRYER2022	1	ISOTROPIC	0	1	THERMAL	0.5	-0.5 ± 2.7	-1.4 ± 1.8	-71 ± 8.8	-60.6 ± 9.3
279	NONE	FRYER2022	5	ISOTROPIC	0	1	THERMAL	0.5	-4.5 ± 3.8	-3.8 ± 2.8	-2.4 ± 5	-2.2 ± 5.5
280	NONE	FRYER2022	10	ISOTROPIC	0	1	THERMAL	0.5	-6.6 ± 4.7	-7.5 ± 3.6	-5.7 ± 5.2	-13.2 ± 5.5
281	PESSIMISTIC	FRYER2012	0.1	ISOTROPIC	0	1	THERMAL	0.5	-21.8 ± 1.8	-25.8 ± 1	-607.3 ± 26.3	-587 ± 26.5
282	PESSIMISTIC	FRYER2012	0.5	ISOTROPIC	0	1	THERMAL	0.5	-19.8 ± 1.9	-22.4 ± 1.2	-607.8 ± 28.9	-585.9 ± 29.1
283	PESSIMISTIC	FRYER2012	1	ISOTROPIC	0	1	THERMAL	0.5	-28.1 ± 2.1	-32.4 ± 1.3	-459 ± 24.6	-441.1 ± 24.8
284	PESSIMISTIC	FRYER2012	5	ISOTROPIC	0	1	THERMAL	0.5	-17.3 ± 3.6	-19 ± 2.6	-56.8 ± 8.7	-49.7 ± 9.1
285	PESSIMISTIC	FRYER2012	10	ISOTROPIC	0	1	THERMAL	0.5	-27.7 ± 4.4	-29.8 ± 3.3	-33.9 ± 6.7	-32.6 ± 7.1
286	PESSIMISTIC	FRYER2022	0.1	ISOTROPIC	0	1	THERMAL	0.5	-28.9 ± 1.8	-33.7 ± 1	-731.4 ± 31.4	-710.4 ± 31.5
287	PESSIMISTIC	FRYER2022	0.5	ISOTROPIC	0	1	THERMAL	0.5	-23.8 ± 1.9	-28.2 ± 1.2	-628.8 ± 29.7	-608.5 ± 29.9
288	PESSIMISTIC	FRYER2022	1	ISOTROPIC	0	1	THERMAL	0.5	-29.1 ± 2.2	-33 ± 1.4	-414 ± 23.5	-396.5 ± 23.7
289	PESSIMISTIC	FRYER2022	5	ISOTROPIC	0	1	THERMAL	0.5	-18.5 ± 3.4	-20.3 ± 2.5	-60.5 ± 8.6	-53.3 ± 9
290	PESSIMISTIC	FRYER2022	10	ISOTROPIC	0	1	THERMAL	0.5	-16.6 ± 4.4	-16.6 ± 3.4	-20.6 ± 6.5	-17.9 ± 6.8
291	OPTIMISTIC	FRYER2012	0.1	ISOTROPIC	0	1	THERMAL	0.5				